GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

January 6, 2004, 01:15:17; Search time 2583 Seconds

(without alignments)

9541.130 Million cell updates/sec

Title:

US-10-088-872-1

Perfect score: 1014

Sequence:

1 atgaaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched:

22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters:

45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100% Listing first 45 summaries

Examples :

EST:*

1: em_estba:*

2: em_esthum:*

3: em estin:*

4: em estmu:*

5: em_estov:*

v: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em gss_vrl:*

28: gb_gss1:* 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

					DOMERCED	,
		ક				
Result		Query		_		Description
No.		Match (Length I	DΒ	ID	Description
		04.0	1552	11	AK076867	AK076867 Mus muscu
1	260.4	84.9	2245	11	AK030474	AK030474 Mus muscu
2	860.4	84.9		11	AK053642	AK053642 Mus muscu
3	360.4	84.9	3039	11	AK076758	AK076758 Mus muscu
4	358.8	84.7	1377	11	AK013205	AK013205 Mus muscu
5	8.44.8	83.3	1449		AU125107	AU125107 AU125107
. 6	770.6	76.0	822	13	BX393735	BX393735 BX393735
. 7	750.8	74.0	1201	11	AK005323	AK005323 Mus muscu
ß	709.4	70.0	1379	10	BG218735	BG218735 RST38476
9	571.8	66.3	784	11	AK013161	AK013161 Mus muscu
10	622.8	61.4	1281		BU116522	BU116522 603139786
<u>1 1</u>	614	60.6	951	13 13	BQ669953	BQ669953 AGENCOURT
12	94.2	58.6	982		BU518807	BU518807 AGENCOURT
13	535.2	57.7	934	13	CD354831	CD354831 UT-M-GMO-
1.4	79.2	57.1	713	14		Bx483012 Homo sapi
1.5	578	57.0	742	2	HSM073180	CA973078 AGENCOURT
2.6	565	55.7	946	1.4	CA973078	BU514920 AGENCOURT
3.7	153.4	54.6	958	13	BU514920	AA278473 %s81h12.x
. 2.3	55 0. 27	54.3	563	9	AA278473	CA982682 AGENCOURT
19	542.6	53.5	930	14	CA982682	CD103801 AGENCOURT
÷0.0	536.2	52.9	732	14	CD103801	BF159587 601769084
21	533.2	52.6	1186	10	BF159587	AW242839 xn26f05.x
େ ଅପର	521.4	51.4	721	9	AW242839	BU914186 AGENCOURT
2.3	517.8	51.1	985	13	BU914186	BQ899617 AGENCOURT
24	500.4	49.3	1060	13	BQ899617	CB559060 AGENCOURT
25	499.6	49.3	817	14	CB559060	BU367571 603787711
26	491.8	48.5	893	13	BU367571	BQ940647 AGENCOURT
27	486.6	48.0	842	13	BQ940647	CB200465 AGENCOURT
2.8	473.6	46.7	854	14	CB200465	BQ021348 UI-H-DH1-
c 29	461	45.5	634	12	BQ021348	CA983082 AGENCOURT
30	460.8	45.4	933	14	CA983082	BI499153 ie27c05.y
31	157 6	45.1	575	12	BI499153	BU905264 AGENCOURT
32	.452	44.6	863	13	BU905264	BI655225 603284039
33	451.4	44.5	927	12		BQ880200 AGENCOURT
34	451.2	44.5		13		BU152376 AGENCOURT
35	445.2	43.9	955	13		AI645170 mt82a02.y
36	445	43.9	534	9	AI645170	AL961440 AL961440
37	443.8	43.8	833	9	AL961440	CB564338 AGENCOURT
c - 38	442.8	43.7	853	14		
39	441.6	43.6	697	12		BI828691 603074707
<i>4</i> ; C	435.8	43.0	1046	14		CB235478 AGENCOURT
z <u>i, J.</u>	435.2	42.9	440	9	AA669484	AA669484 af74g07.x
c 42	434.2	42.8		13		BU631151 UI-H-FE1-
43	433	42.7	961	13		BU129770 603117261
4.4	430	42.4	752	14		CB524885 UI-M-FY0-
45	428	42.2	985	14	CA970822	CA970822 AGENCOURT

ALIGNMENTS

```
RESULT 1
AK076867
                                                                    HTC 07-DEC-2002
                                      1552 bp
                                                  mRNA
                                                          linear
LCCUS
            AK076867
            Mus musculus adult male testis cDNA, RIKEN full-length enriched
DEFINITION
            library, clone:4930520C08 product:MO25-LIKE PROTEIN homolog [Homo
            sapiens], full insert sequence.
            AK076867
ACCESSION
VERSION
            AK076867.1 GI:26345723
KEYWORDS
            HTC; CAP trapper.
SOURCE
            Mus musculus (house mouse)
  ORGANISM
            Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
            Carninci, P. and Hayashizaki, Y.
  AUTHORS
            High-efficiency full-length cDNA cloning
  CITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
            99279253
  MEDLINE
            10349636
   PUBMED
            2
ABFERENCE
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  MUTHORS
             Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Mormalization and subtraction of cap-trapper-selected cDNASE to
  TO SEE
             prepare full-length cDNA libraries for rapid discovery of new genes.
            Genome Res. 10 (10), 1617-1630 (2000)
                                                                         1. 2
  JOURNAL
                                                                         145
            20499374
  MEDLINE
            11042159
  DUBMED
REFERENCE
            Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  NUTHORS
            Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Iton, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
            sequencing pipeline with 384 multicapillary sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
            20530913
  MEDLINE
            11076861
  PUBMED
REPERENCE
            Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
AUTHORS
            Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
            Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
            Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
            Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
             Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
            Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
            Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
            Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
            Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
```

Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,

```
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
            Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
            Marchionni, I.., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
            Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
            Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
            Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
            Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
            and Hayashizaki,Y.
            Functional annotation of a full-length mouse cDNA collection
  TITLE
            Nature 409 (6821), 685-690 (2001)
  JOURNAL
            21085660
  MEDLINE
            11217851
   PUBMED
REFERENCE
            The FANTOM Consortium and the RIKEN Genome Exploration Research
  AUTHORS
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
  TITLE
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
  JOURNAL
                (bases 1 to 1552)
REFERENCE
            Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
  AUTHORS
            Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
            Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
            Miraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
            Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
            Coya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
            Nemura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
            Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
            Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
            Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
             Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
             Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
             Direct Submission
  TITLE
             Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
  JOURNA ...
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             WRL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             %ax:81-45-503-9216)
             cDNA library was prepared and sequenced in Mouse Genome
 COMMENT
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
                       Location/Qualifiers
 FEATURES
                       1. .1552
      source
                       /organism="Mus musculus"
                       /mol_type="mRNA"
                       /strain="C57BL/6J"
                       /db_xref="FANTOM_DB:4930520C03"
                       /db_xref="MGI:1894876"
                       /db_xref="taxon:10090"
                       /clone="4930520C03"
```

/sex="male"

```
/tissue_type="testis"
               /clone_lib="RIKEN full-length enriched mouse cDNA library"
               /dev stage="adult"
               316. .1320
   CDS
               /note="unnamed protein product; MO25-LIKE PROTEIN homolog
               [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
               98.2%ID, 100%length, match=1002)
               putative"
               /codon start=1
               /protein_id="BAC36513.1"
               /db xref="GI:26345724"
               /db xref="MGI:1914081"
               translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ/
               AMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL
               RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
               QFRDFFKYVELSTFDTASDAFATFKDLLTRHKVLVADFLEQNYDTTFEDYEKLLQSEN
               YVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
               {	t FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP}
               1539. .1544
   polvA signal
               /note="putative"
               1552
   polyA_site
               /note="putative"
                                401 t
           490 a
                  320 G
                         341 a
PASE COUNT
ORIGIN
                                         Length 1552;
                         Score 860.4; DB 11;
                   84.9%;
 Query Match
                         Pred. No. 5.8e-176;
 Best Local Similarity
                   90.5%;
                                        Indels
                                                   'Gaps.
                           Mismatches
                                     96;
                        0 ;
 Matches 918; Conservative
        1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
\mathbb{Q}\mathcal{Y}
          307 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAAATC 366
D
        ()x_i
                    11111111
       Db.
       QΣ
          181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
QУ
          437 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 546
dC
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Qy
          547 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 606
Db:
       361 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
ĆΛ
          607 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 666
Db
       361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
QУ
           667 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 726
Db
```

£1	C1 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
	27 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 786
Qy 4.	31 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGIGGAGTTGTCAACATTTGATATIGCT 540
Db 71	37 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 846
Qy 5	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
Db 84	
Qy 6	1 TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
. Db 90	
Qy 6	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Db 9	
Qy 7:	21 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
Db 13	
7	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840:
Бил. (140)	
3	1 UTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
da da	GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1206
3 3.	01 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Db 12	
Qy 9	61 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Db 12	
RESULT 0 AK030474	
Locus	AK030474 2245 bρ mRNA linear HTC 05-DEC-2002
Definition	Mus musculus adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:5330416K15 product:MO25-LIKE PROTEIN
a depositor	homolog [Homo sapiens], full insert sequence.
ACCESSION VERSION	AK030474 GI:26326468
KEYWORES SOURCE	HTC; CAP trapper. Mus musculus (house mouse)
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE AUTHORS	1 Carninci, P. and Hayashizaki, Y.

```
High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
  MEDLINE
            99279253
            10349636
   PUBMED
            2
REFERENCE
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
            20499374
  MEDLINE
            11042159
   PUBMED
REFERENCE
            Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
             Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
             Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
             Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
             Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
             Yoneda, Y. Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
             Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
             RIKEN integrated sequence analysis (RISA) system -- 384-format
  TITLE
             sequencing pipeline with 384 multicapillary sequencer
             Genome Res. 10 (11), 1757-1771 (2000)
  COURNAL
             20530913
  MEDLINE
   DUBMED
             11076861
PARENCE
             Mawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., 🔀
  AUTHORS
             Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., A
             Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
             Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
             Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
             Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
             Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
             Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
             Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
             Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
             Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
             Pletcher,C., Fujita,M.; Gariboldi,M., Gustincich,S., Hill,D.,
             Wofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
             Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
             Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
             Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
             Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
             Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
              and Hayashizaki,Y.
             Functional annotation of a full-length mouse cDNA collection
   THILE
              Mature 409 (6821), 685-690 (2001)
   JOURNAL
   MEDLINE
              21085660
              11217851
    PURMED
 REFERENCE
              The FANTOM Consortium and the RIKEN Genome Exploration Research-
   AUTHORS
              Group Phase I & II Team.
              Analysis of the mouse transcriptome based on functional annotation
   TITLE
              of 60,770 full-length cDNAs
              Nature 420, 563-573 (2002)
   JOURNAL
                 (bases 1 to 2245)
 REFERENCE
              Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
```

AUTHORS

```
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
      Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
      Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
      Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
      Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
      Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
      Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
      Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
      Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
      Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
      Muramatsu, M. and Hayashizaki, Y.
      Direct Submission
      Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
       Physical and Chemical Research (RIKEN), Laboratory for Genome
       Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
       PIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
       Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
       URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
       Fax:81-45-503-9216)
       cDNA library was prepared and sequenced in Mouse Genome
       Encyclopedia Project of Genome Exploration Research Group in Riken
       Genomic Sciences Center and Genome Science Laboratory in RIKEN.
       Division of Experimental Animal Research in Riken contributed to
       prepare mouse tissues.
       Please visit our web site for further details.
       URL:http://genome.gsc.riken.go.jp/
       URL:http://fantom.gsc.riken.go.jp/.
                Location/Qualifiers
                1. .2245
ವರಿಗಳು ಡಿಕ
                /organism="Mus musculus"
                /mol type-"mRNA"
                /strain="C57BL/6J"
                /db xref="FANTOM DB:5330416K15"
                /db xref="taxon:10090"
                /clone="5330416K15"
                /sex="male"
                /tissue_type="pituitary gland"
                /clone lib="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="adult"
                300. .1313
                /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                 [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                98.2%ID, 100%length, match=1002)
                putative"
                /codon start=1
                /protein_id="BAC26978.1"
                /db xref="GI:26326469"
                translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK/
                SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKEPTDDEQFADEKNYLIKQIRDLKK
                AAP"
```

581 t

480 c

667 a

517 g

TITLE.

COMMENT

RES

CDS

BASE COUNT

ORIGIN

JOURNA!.

Query Ma Best Loc Matches	cal S	84.9%; Score 860.4; DB 11; Bength 2243; Similarity 90.5%; Pred. No. 6.2e-176; B; Conservative 0; Mismatches 96; Indels 0; Gaps	0. ;
Qу		ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 6	
Db			
Qy		CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	
Db	360	CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	119
Qλ.		GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	
Db	420	GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAAC	179
Qy		CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG	
Db		CCCCTACAGAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG	
QY		ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA :	
Db -			
Qy	500	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT :	
อิธ			
1		GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	
DS	1.0	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	
	,		
ОУ		TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	
Db			
Qy		TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	
වර්			
Э.У.		TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	
Dр) TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG	
:		AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	
Dρ		© AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC	
δλ		AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	
Db	10,2	0 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC	
	пĠ	CTCCTTCCCCATAAAACTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840

震:

```
1030 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1139
Db
         841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
QУ
             1140 GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1199
Db
       901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
QΥ
             1200 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGACGACGACGAGCAGCTTTGCTGAC 1259
Db
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QУ.
             1260 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1313
DD
RESULT 3
AK053642
                                                            HTC 05-DEC-2002
                                            mRNA .
                                                    linear
                                  3039 bp.
           AK053642
LOCUS
          Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
DEFINITION
           library, clone:E130116021 product:M025-LIKE PROTEIN homolog [Homo
           sapiens], full insert sequence.
ACCESSION
           AK053642
           AK053642.1 GI:26343600
VERSION
           HTC; CAP trapper.
KEYWORDS
           Mus musculus (house mouse)
COURCE
  CRGANIES Mus musculus
           Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
           Carninci, P. and Hayashizaki, Y.
  AUTHORS
           High-efficiency full-length cDNA cloning
  CIPLE
           Meth. Enzymol. 303, 19-44 (1999)
                                                                 166
  ZANSU TENAL
           99279253
                                                                 W 1967
  MEDDLINE
           10349636
   3 JBMED -
REFERENCE
           Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  MUTHORS
           Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
           Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE.
           prepare full-length cDNA libraries for rapid discovery of new genes
           Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
           20499374
  MEDLINE
           11042159
   FUBMED.
REFERENCE
           Shibata, K., Iton, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
           Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
           Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
           Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
           Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
           Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
           Okazaki.Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
           RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
            sequencing pipeline with 384 multicapillary sequencer
           Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
           20530913
  MEDITINE
            11076861
   PUBMED
REFERENCE
```

```
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 AUTHORS
           Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
           Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
           Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
           Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
           Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
           Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
           Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
           Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
           Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
           Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
           Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
            Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
           Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
            Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
            Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
            Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
            Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
            and Hayashizaki, Y.
            Functional annotation of a full-length mouse cDNA collection.
  TITLE
            Nature 409 (6821), 685-690 (2001)
  JOURNAL
            21085660
 MEDLINE
            11217851
  PUBMED
REFERENCE
            The FANTOM Consortium and the RIKEN Genome Exploration Research
  AUTHORS
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
  THIE
            of 50,770 full-length cDNAs
            Mature 420, 563-573, (2002)
  JOURNAL
               (bases 1 to 3039)
            Adachi, J., Aizawa, K., 'Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
REFERENCE
  AUTHORS
            Fukuda.S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
            Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
            Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
            Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
            Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
             Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
            Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
             Muramatsu, M. and Hayashizaki, Y.
             Direct Submission
  TITLE
             Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
  JOURNAL
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
             cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
COMMENT
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
```

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

```
URL: http://fantom.gsc.riken.go.jp/.
               Location/Qualifiers
FEATURES
               1. .3039
   source
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="C57BL/6J"
               /db xref="FANTOM DB:E130116021"
               /db xref="taxon:10090"
                /clone="E130116021"
                /tissue_type="eyeball"
               /clone_lib="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="0 day neonate"
               287. .1300
   CDS
                /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                98.2%ID, 100%length, match=1002)
                putative"
                /codon start=1
                /protein_id="BAC35457.1"
                /db_xref="GI:26343601"
                /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
                AAP"
                3022. .3027
    polyA_signal
                /note="putative"
                3039
    rolyA_site
                /note="putative"
                          676 a
                                 336 t
                   645 c
BASE COUNT
            882 a
ORIGIN
                          Score 860.4; DB 11: Length 3039;
                    84.98;
  Quary Match
                          Pred. No. 6.6e-176;
                    90.5%;
  Bost Local Similarity
                                          Indels
                                                    Gaps
                         0; Mismatches
                                      96;
  Matches 918; Conservative
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
ŨΥ
           237 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 346-
Db
        Qy
           Eh
        Qу
           Db
        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Çy
           467 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 526
Db
        241 ACACTGATAGCTGACCTGCACCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300.
 Qу
           527 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 586
 Db
```

ĎΆ	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAG'	i
Ob	587 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCATCAC	
Ωу	363 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTT.	1
Οb	THE TRANSPORT OF THE TR	
Qy Db		'A. 766
Qy Db	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGC	i
Qy Db .	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGA 	
Qy Db	601. TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGA 	1 1
27 72	361 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCA 	1 1
Qy 	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGA 	1 1
dy Es	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGT	TT 840?
0y Db	GET GTGGCCAGTCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCA	AA 90:0:
	901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTG	AC 960.
©√ Db	961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 101 	

RESULT 4 AK076758 linear 13'77 bp mRNA LCCU3 DEFINITION

HTC 07-DEC-2002 Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930433N18 product:MO25-LIKE PROTEIN homolog [Homo sapiens], full insert sequence.

AK076758 ACCESSION

```
AK076758.1 GI:26345637
VERSION
            HTC; CAP trapper.
KEYWORDS
            Mus musculus (house mouse)
SOURCE
            Mus musculus
  ORGANISM
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
            Carninci, P. and Hayashizaki, Y.
  AUTHORS
            High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
             99279253
  MEDLINE
             10349636
   DIJBMED
REFERENCE
             2
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
             Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
             prepare full-length cDNA libraries for rapid discovery of new genes
             Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
  MEDLINE
             20499374
             11042159
   PUBMED
REFERENCE
             Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
             Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
             Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
             Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
             ⊮ujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
             Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
             Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
  图1:1.图
             sequencing pipeline with 384 multicapillary sequencer
             Genome Res. 10 (11), 1757-1771 (2000)
  CCRNAL
             20530913
  EDLINE
   POBMED
             11076861
  DEFENCE
             Mawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
  MOTHORS:
             Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
             Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
             Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
             Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
             Pleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
             Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
             Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
             Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
             Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
             Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
             Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
             Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
             Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
             Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
             Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
             Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
             Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
             and Hayashizaki, Y.
             Functional annotation of a full-length mouse cDNA collection
   TITLE
             Nature 409 (6821), 685-690 (2001)
   JOURNAL:
              21085660
   MEDLINE
```

11217851

PUBMED

```
The FANTOM Consortium and the RIKEN Genome Exploration Research
REFERENCE
  AUTHORS
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
  TITLE
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
  JOURNAL
               (bases 1 to 1377)
            Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
REFERENCE
            Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
  AUTHORS
            Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
            Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
            Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
            Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
            Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
            Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
            Shibata K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
            Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
            Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
            Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
             Direct Submission
             Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
  TITLE
  JOURNAL
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             PIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Ranagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             TRL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
                                                                                   F.
             max:81-45-503-9216)
              ONA library was prepared and sequenced in Mouse Genome
                                                                                   c!
             Encyclopedia Project of Genome Exploration Research Group in Riken e
 COMMENT
             denomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL: http://fantom.gsc.riken.go.jp/.
                       Location/Qualifiers
 PEATURES
                       1. .1377
      source
                       /organism="Mus musculus"
                       /mol type="mRNA"
                       /strain="C57BL/6J"
                       /db_xref="FANTOM_DB:4930433N18"
                       /db_xref="MGI:1894867"
                       /db xref="taxon:10090"
                       /clone="4930433N18"
                       /sex="male"
                       /tissue_type="testis"
                       /clone_lib="RIKEN full-length enriched mouse cDNA library"
                       /dev_stage="adult"
                       287. .1300
                       /note="unnamed protein product; MO25-LIKE PROTEIN homolog
       CDS
                       [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                       98.2%ID, 100%length, match=1002)
                       putative"
                       /codon_start=1
                       /protein_id="BAC36470.1"
                       /db xref="GI:26345638"
                        /db xref="MGI:1914081"
```

/translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK AAP"

BASE COUNT ORIGIN 430 a 294 c 306 g 347 t

Query M	atch		84.7%;	Score	858.8;	DB 11;	Length	1377;		
Best Lo Matches	cal S 917	imilarity ; Conserva	90.4%; tive	Pred.	No. 1.3e smatches	97;	Indels	0;	Gaps	0;
Qy .		ATGAAAAAAAT	+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1		1 1 1 1 1	11111	1111111	1		
Db		ATGAAAAAAAT	GCCTTTG7	TTAGTA	AATCACACA	AAAATC	CAGCAGAA	ATTGTC		
Qý		CTGAAAGACAA			! !		11111	! i ! ! ! !	7 1 1 1	
aG		CTGAAAGACAA								
ЭУ		GAAGTGTCTAA		1 1 1 1		1	1111	11111		
Ti.							;			
		CCCCCAACAGA	111111	11111		11 111		1		
		ACACTGATAGO								
		ACACTGATAGO	9411111			111111	4 1 1 1 1	11111	. 1 !	
		TTTAACAACAT								
Dp Dp		TTCAACAACAT	. i . i i i i i i		1.11			1111		
ĜÃ.		GCTCATCCTCA		111111				. 1 I i I I	11111	
Dp.		TCTCATCCTC								
Qy	1.5	CGTTGTGGGA'	11 i i l l l i	+1+1+1		1111				
Db	1.5	CGCTGTGGGA'		:						
QY	44.14	TTTTCTAATC	1 111111	1111111		11 111			111 111	
Db ⁻	767	TTTTCTAATC.	AGTTCAGA	GATTTCT	TTCAAGTAT	GTTGAG	CTGTCCAC	CTTTGA	TATCGCT	826
ØУ	-541	TCAGATGCCT	1 111111				111111	- 1 1 1	111111	
Db	827	TCAGATGCCT	TCGCTACT							
Qy	·	TTCTTAGAAC	111111		1	11111		11111	_	
Db (887	TTCTTAGAAC	AAAATTA	rgacact:	ATTTTTGA	AGACTAT	GAGAAACT	GCTGCA	ATCTGAG	946

```
661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
QУ
            947 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1006
Db
        721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
ÇУ
            1007 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1066
Db
        781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
QУ
            1067 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1126
Db
        341 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
0.0
            1127 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAAATCAGCCCAAA 1186
\Gamman
        901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Qy
            1187 CTCATTGAGTTTTTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1246
Db.
        961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QУ
            1247 GAGAAGAACTACC'IGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1300
HASULT S
EK013205
                                        mKNA
                                                        HTC 05-DEC-2002
                               1449 bp
                                                linear
          AK013205
LOCUS
          Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DEFINITION
          enriched library, clone:2810430N08 product:MO25-LIKE PROTEIN
          homolog [Homo sapiens], full insert sequence.
ACCESSION
          AK013205
          AK013205.1 GI:12850419
VERSION
          HTC; CAP trapper.
KEYWORDS
          Mus musculus (house mouse)
GOURCE
          Mus musculus
  ORGANISM
          Eukaryota; Metazoa; Chordatá; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae: Mus.
REFERENCE
          Carninci, P. and Hayashizaki, Y.
  AUTHORS
          High-efficiency full-length cDNA cloning
  TITLE
          Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
           99279253
  MEDLINE
           10349636
   PUBMED
REFERENCE
           Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
           Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
           Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
           prepare full-length cDNA libraries for rapid discovery of new genes
           Genome Res. 10 (10), 1617-1630 (2000)
  JCURNAL
  MEDLINE
           20499374
  PURMED
           11042159
REFERENCE
           Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
           Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
```

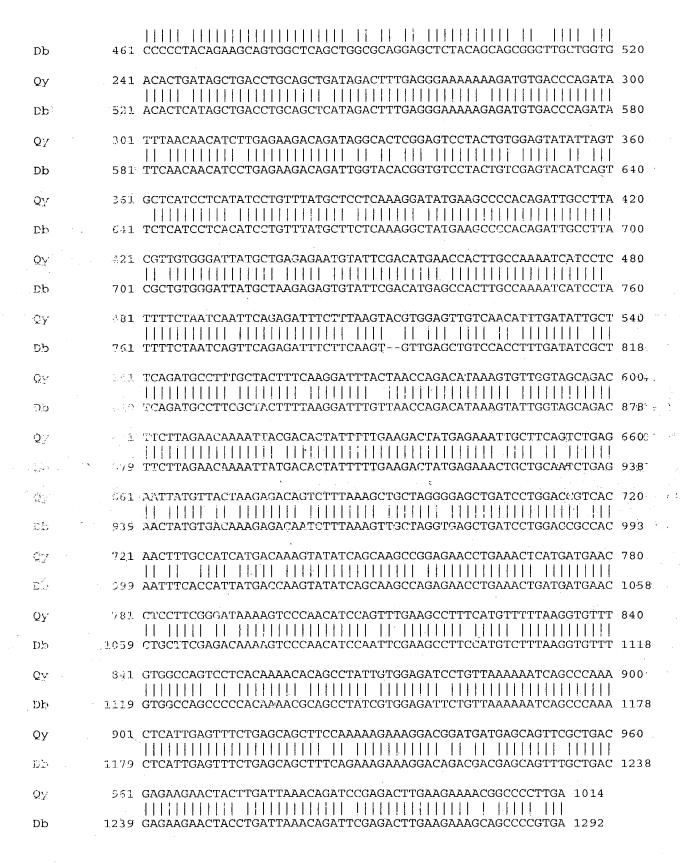
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

```
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Ycneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
 TITLE
            sequencing pipeline with 384 multicapillary sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
            20530913
  MEDLINE
            11076861
   PUBMED
REFERENCE
            Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
  AUTHORS
            Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
            Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
            Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
            Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
            Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
            Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
            Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
            Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
            Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
             Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
             Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
             Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
             Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
             Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
             Cabo, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
             Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
             Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
             and Hayashizaki,Y.
             Junctional annotation of a full-length mouse cDNA collection
  TITLE
             Nature 409 (6821), 685-690 (2001)
  JOURNAL
             21085660
  MEDLINE
             11217851
   PUBMED
REFERENCE
             The FANTOM Consortium and the RIKEN Genome Exploration Research
   AUTHORS
             Group Phase I & II Team.
             Analysis of the mouse transcriptome based on functional annotation
   TITLE
             of 60,770 full-length cDNAs
             Nature 420, 563-573 (2002)
   JOURNAL
             6 (bases 1 to 1449)
REFERENCE
             Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
   AUTHORS
             Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
             Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
             Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
              Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
              Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
              Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Cwa, C.,
              Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
              Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
              Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
              Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
              Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
              Direct Submission
   TITLE
              Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
   ÖURNAL
              Physical and Chemical Research (RIKEN), Laboratory for Genome
              Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
              RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
```

4

```
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
         URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
         Fax:81-45-503-9216)
         Please visit our web site (http://genome.gsc.riken.go.jp/) for
COMMENT
         further details.
         eDNA library was prepared and sequenced in Mouse Genome
         Encyclopedia Project of Genome Exploration Research Group in Riken
         Genomic Sciences Center and Genome Science Laboratory in RIKEN.
         Division of Experimental Animal Research in Riken contributed to
         prepare mouse tissues. First strand cDNA was primed with a primer
         prepared by using trehalose thermo-activated reverse transcriptase
         and subsequently enriched for full-length by cap-trapper. cDNA went
         through one round of normalization to Rot = 7.5 and subtraction to
         Rot = 37.5. Second strand cDNA was prepared with the primer adapter
         of sequence [5'
         with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
         Host: SOLR.
                Location/Qualifiers
PEATURES
                1. .1449
    source
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db xref="FANTOM_DB:2810430N08"
                /db_xref="MGI:1893512"
                /dh xref="taxon:10090"
                /cloue="2310430N08"
                /tissue_type="whole body"
                /clone_lib="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="10, 11 days embryo"
                281. .1292
       4 ture
                /note="MO25-LIKE PRCTEIN homolog [Homo sapiens]
                (SWISSPROT|Q9H9S4, evidence: FASTY, 98.2%ID, 100%length,
                match=1002)
                putative"
                /db xref="MGI:1914081"
                                 367 t
                    304 c
                          325 g
             453 a.
radár cousta
ONEG.III
                    83.3%; Score 844.8; DB 11;
                                           Length 1449;
  Query Match
                    90.2%; Pred. No. 1.4e-172;
  Best Local Similarity
                          0; Mismatches
                                                     Gaps : 1;
                                       97; Indels
                                                  2:
 Matches 915; Conservative
          1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
 QÝ:
           281 ATGAAAAAATGCCCTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 340
Ľb
         СŸ
           D_{2}^{1}
        QY
            Db
        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
```

QУ



```
AU125107
                                                           EST 01-AUG-2002
                                            mRNA
                                                   linear
                                  822 bp
LOCUS
           AU125107 NT2RM4 Homo sapiens cDNA clone NT2RM4001047 5', mRNA
DEFINITION
           sequence.
ACCESSION
           AU125107
           AU125107.1 GI:10949823
VERSION
           EST.
KEYWORDS
           Homo sapiens (human)
SOURCE
           Homo sapiens
  ORGANISM
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 822)
REFERENCE
           Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J.,
  AUTHORS
           Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and
           Isoqai, T.
           HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S.,
  TITLE
           Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki
            Y., Sugano, S., Isogai, T.)
           Unpublished
  JOURNAL
           Contact: Takao Isogai
~COMMENT
           Genomics Laboratory
           Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            mail: genomics@hri.co.jp
            TRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Melix Research Institute.
                    Location/Qualifiers
LEATURES
                    1. .822
     sowices.
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    /clone="NT2RM4001047"
                    /cell_type="teratocarcinoma"
                    /cell_line="NT2"
                    /clone lib="NT2RM4"
                    /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
                    precursor cells"
                               171 g 216 t
                                                   3 others
                        164 C.
 PASE COUNT
               268 a
 ORIGIN
                         76.0%; Score 770.6; DB 9; Length 822;
  Query Match
                         98.5%; Pred. No. 1.5e-156;
  Best Local Similarity
  Matches 798; Conservative 0; Mismatches
                                              10;
                                                    Indels
           19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
 ÇУ
              1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
 T.b.
           79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
 Ç.y.
              61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120
 Db
          Çу
```

	1.1		
Db	121 CA	AAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	
Qу	1.1	CTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258	
Dβ	181 GC	CTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG	
QУ		AGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318	
Db	241 C	AGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 900	
ÕΆ	1	GACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378	
Db	301 A	GACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATMTCCTC	
Qy		TTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438 	
Db			
SA.		GAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 49 	
ф		SATTTCTTTAAGTACGIGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 55	
Qy :	1	GATTTCTTTAAGTACGIGGAGIIGICAACATIIGAAAAAAAAAA	
5	::::, ·	TCAAGGATITACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 61	:8 50
Ω Υ `. υ(δ)	and the second s	TTCAAGGATTTACTAACCAGACATAAAGTOTTGG MGGAGACTTCTTAGAACAAAATTAC 60 	
Ç.	. 1	BACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 57	
D).is	501 9	ACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTTTGTTTTGAGAAATTGCTTCAGAAATTGCTTCAGAAATTGCTTCAGAAATTGCTTCAGAAATTGCTTCAGAAATTGCTTCAGAAATTGCTTCAGAAATTGCTTCAGAAATTGCTTCAGAAATTTGCTTCAGAAATTGCTTCAGAAATTTTGTTTTGTTTTTGTTTTTGTTTTTTTT	
QУ	,	CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 73	
Dр	, €51 C	CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTIGCCATCATGACA	
δλ	1	AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 79	,
Σb		AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 7	
ΟY	15.8	CCCAACATCCAGTTTGAAGCCTTTCATGTT 828	
1)b	<i>739</i> (CCCAACATCCAGTTTGAGCCTTCTGGTTTT 808	
BX	SULT 7 393735	1201 bp mRNA linear EST 13-MAY-2	003
	VC NATULEOUS ON	1201 bp mkNA IIHear EST IS IN 2 393735 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapie NNA clone CSODC002YI01 5-PRIME, mRNA sequence.	ns
	CESSION BX	393735 393735.1 GI:30624044	
KE	YWORDS ES	er.	
SC	URCE Ho	omo sapiens (human)	

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ORGANISM Homo sapiens
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            (bases 1 to 1201)
REFERENCE
         Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
         Full-length cDNA libraries and normalization
 AUTHORS
 TITLE
         Unpublished
 JOURNAL.
          Contact: Genoscope
COMMENT
         Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          Library was constructed by Life Technologies, a division of
          Invitrogen. This sequence belongs to sequence cluster 6951.r For
          more information about this cluster, see
          http://www.genoscope.cns.fr/
          cgi-bin/cluster.cgi?seq=CS0DC002AE01QP1&cluster=6951.r. Contact :
          Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CS0DC002AE01QP1.
                 Location/Qualifiers
PEATURES
                 1. .1201
    source
                 /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db xref="taxon:9606"
                  /clone="CS0DC002YI01"
                  /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                  /clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                  /note="1st strand cDNA was primed with a NotI-oligo(dT)
                  primer. Five prime end enriched, double-strand cDNA was
                  digested with Not I and cloned into the Not I and EcoR V
                  sites of the pCMVSPORT 6 vector. Library was normalized."
                                            70 others
                            239 g 321 t
                     223 C
              348 a
 BASE COUNT
 CREGIN
                            Score 750.8; DB 13; Length 1201;
                      74.0%;
  Query Match
                            Pred. No. 3.1e-152;
                      91.0%;
  Post Local Similarity
                                                        Gaps
                                              Indels
                           45; Mismatches
                                          27;
  Matches 766; Conservative
           1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
            ÕΑ
         323 ATGAAAAAATGCCTTTGTTTAGTWAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 382
 Db:
          QУ
         Db
         QУ
          443 GWAGTGTCTWAWTCACTGCTAGCWATGWWAGATATTYTGTGTGGTACAWACGAGWAAGAT 502
 Db
          181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
             QY
          503 CCCCCAACAGAAGCTGTGGCTCAGCTTGCAYAAGWTYTCTTCAGYWGTTGCCTGCTAGTG 562
  Db
          241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
             ĊУ
          563 ACACTGATWGCTGACCTGCAGCTGATAGACTTTKAGGGAAAARDAGATGTGACCCAGATT 622
  áŒ
```

	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATA	TTAGT	360
Qу		11111	
Db	623 TTTTACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATA	11 1110 -	
Qy.	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTG	1 1 1 1 1	
Db	583 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTT		
Qy	421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCA	ATCCTC	480:
Db	421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACGACTTGCGACATGAACGACTTGCGACATGAACGACTTGCCGAAAATCACTGACATGTGTGGGATTATGCTGAGAAGAATGTATTCGACATGAACCACTTGCCAAAAATCACTGACATGAACCACTTGCCAAAAATCACTGACATGACATGAACCACTTGCCAAAAATCACTGACATGAACCACTTGCCAAAAATCACTGACATGAACCACTTGCCAAAAATCACTGACATGAACCACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGCCAAAAATCACTGACATGAACGACTTGACAAAGATGAACGACTGAACGACTGAACGACTGAACGACTGAACGACTGAACGACTGAACGACTGAACGACTGAACGACTGAACGACACTGAACGACTGAACGACTGAACGACATGAACGACTGAACGACTGAACGACATGAACGACTGAACGACATGAACGACTGAACGACATGAACGACATGAACGACTGAACGACACTGAACGACACTGAACGACATGAACGACATGAACGACATGAACGACATGAACGACACTGAACGACACTGAACGACACTGAACGACACTGAACGACACTGAACGACACTGAACACACAC	ATCCTC	802
	THE A THE A CACATTTCTTTA A GTACGTGGAGTTGTCAACATTTGATA	ATTGCT	
QУ	481 TTTTCTAATCAATTCAGAGATTTCTTTAGGTACTTCTTTAGTAGTACTTCTTTTTTTT		
Db	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTA		
Ωу	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAOTATTTTCAAGGATTTACTAACCAOTATTTTTTTTTT		
ינעם.	COLOR CONTRACTOR CONTRACTOR A CALCON TO CALCON TO CONTRACTOR CONTRACTOR A CALCON TO CA	TCTGAG	
Qy Db	601 TTCTTAGAACAAAATTACGACACTATTTTTGAACACTATTTTGAACACTATTTTTGAACACTATTTTTGAACACTATTTTTTTT	1 1 1 1 1 1	
	661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTA-GGGGAGCTGATCCTGGA	CCGTCA	719
. b	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTA GGGGAGGTGATCCTGGA 	1:11:	
	TO TO TO TO TO TO TO TO TOTATO TO TOTATO TO THE TOTATO THE TOTATO TO THE TOTATO T	GATGAA	
Qy Db	CAACTTTGCCATCATGACAAAGTWTATYMYCYMBCCGGSBYHCCYSWWACTCMT	rgatgaa	11022
Qy	780 CCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAA	• 1 1 1 1	
Pa du	: : :	€KGTTTT	1159
СА	340 TG 841) = 1
· Db	1160 TG 1161		•
RESULT		i i	
AKO0532 LOCUS DEFINIT	AK005323 1 1379 bp middl	05-DE0	riched
ioe e an a a	ION Mus musculus adult male cerebellum CDNA, KIRAN 2020 Library, clone:1500031K13 product:MO25-LIKE PROTEIN hom sapiens], full insert sequence.	olog (i	HOIIIO
ACCESSI	ON AK005323	er er	
VERSION	AK005323.1 GI:12837793		
KEYWORI SOURCE	Mus musculus (house mouse)		e .
ORGAI	ISM Mus musculus	eleost	omi;
	Eukaryota; Metazoa; Chordata; Cranrata, Volumerata; Muridae; Mu Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu	ırinae;	Mus.
REFEREI	CE 1		
ACTIN			

```
High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
            99279253
  MEDLINE
            10349636
   PUBMED
            2
REFERENCE
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
             20499374
  MEDLINE
             11042159
   PUBMED
REFERENCE
             Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
             Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
             Sumi, N. Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
             Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
             Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
             Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
             Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
             sequencing pipeline with 384 multicapillary sequencer
             Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL .
             20530913
  MEDLINE
             11076861
   PUBMED
  DENCE
             Mawai, J. Shinagawa, A., Shibata, K., Yoshino, M., Iton, M., Ashii, Y.,
  LECTIONS
             Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
             Alzawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
             Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito R.,
             Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
             Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
             Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
             Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
             Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
             Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
             Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
             Pleacher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
             Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
             Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
             Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
             Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
             Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
             Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
             and Hayashizaki, Y.
             Functional annotation of a full-length mouse cDNA collection
   TITLE .
             Nature 409 (6821), 685-690 (2001)
   JOURNA'
             21085660
   MEDLINE
             11217851
    PUBMED
 REFERENCE
             The FANTOM Consortium and the RIKEN Genome Exploration Research
   AUTHORS
             Group Phase I & II Team.
             Analysis of the mouse transcriptome based on functional annotation
   TITLE
             of 60,770 full-length cDNAs
             Nature 420, 563-573 (2002)
   JOURNAL
                 (bases 1 to 1379)
 REFERENCE
             Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
   AUTHORS
```

```
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
     Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
     Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
     Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
     Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
     Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
     Saitc, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
     Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
     Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
     Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
     Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
     Direct Submission
     Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
     Physical and Chemical Research (RIKEN), Laboratory for Genome
     Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
     RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
     Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
     URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
      Fax:81-45-503-9216)
      Please visit our web site (http://genome.gsc.riken.go.jp/) for
      further details.
      cDNA library was prepared and sequenced in Mouse Genome
      Encyclopedia Project of Genome Exploration Research Group in Riken
      Genomic Sciences Center and Genome Science Laboratory in RIKEN.
      Division of Experimental Animal Research in Riken contributed to
      repare mouse tissues. First strand cDNA was primed with a primer
      greepared by using trehalose thermo-activated reverse transcriptase \mathbf{p}_{AB}
      and subsequently enriched for full-length by cap-trapper. Second
      strand cDNA was prepared with the primer adapter of sequence[5'
      was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3'
      end: SstI. Host: SOLR.
               Location/Qualifiers
               1. .1379
Dourse
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="C57BL/6J"
               /db_xref="FANTOM_DB:1500031K13"
               /db xref="MGI:1901050"
               /db xref="taxon:10090"
               /clone="1500031K13"
               /sex="male"
               /tissue_type="cerebellum"
               /clone_lib="RIKEN full-length enriched mouse cDNA library"
               /dev_stage="adult"
               285. .1175
               /note="unnamed protein product; MO25-LIKE PROTEIN homolog
               [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
               98.2%ID, 100%length, match=1002)
               putative"
               /codon_start=1
               /protein_id="BAB23953.2"
               /db_xref="GI:26342524"
               /db xref="MGI:1916258"
```

PLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN

TITLE

COMMENT

THURES

JCURNAL

NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL ${\tt FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT1FEDYEKLLQ}$ SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV

FKNSVFITNRIHGLKRWLSS" 1361. .1366 polyA_signal /note="putative" polyA_site /note="putative" 352 t 284 C 291 q

452 a BASE COUNT ORIGIN

Length 1379; Score 709.4; DB 11; 70.0%; Query Match 88.9%; Pred. No. 3e-143; Best Local Similarity Gaps Indels 0; Mismatches 96; 0; Matches 767; Conservative 1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAAATTGTGAAAAATC 60 Qy 285 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 344 Db Çу Db QУ \supset b 191 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAACTCTACAGCAGTGGCCTGCTAGTG 240 : T 455 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 524 241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300 $O_{\mathcal{F}}$ S25 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 584 Tib) 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360 W 585 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 644 Db 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420 QΥ 545 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 704 Db 421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480 O_{λ} 705 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 764 Db. 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540 Qy. TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 824 D_{ij}^{lo} $561: {\tt TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC} {\tt 500}_{1:1}$ QY825 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 884 Db 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660 Qy

```
885 TTCTTAGAACAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 944
Db
         661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
             QУ
         945 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1004
Db
         721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
             QУ
        1005 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGAAC 1064
Db
         781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840 .
             Qу
        1065 CTGCTTCGAGACAAAGTCCCAACATCGAAGTCCTATCGAAGCCTTCCATGTCTTTAAGAATTCT 1124
Db
          841 GTGGCCAGTCCTCACAAAACACA 863
QУ
                1125 GTCTTTATTACAAATAGAATACA 1147
Do
RESULT 9
DG218735
                                                            EST 21-APR-2001
                                                    linear
                                  784 bp mRNA
           BG218735
           RST38476 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
LOCUS
DEFINITION
           3G218735
ACCESSION
            BG218735.1 GI:13744756
 TERSION
            EST.
 KUTTWORDS
            Homo sapiens (human)
 CURCE
                                                                 1346-
           Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   ANISM
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            : (bases 1 to 784)
           Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
 CRENCE
            Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.
  AUTHORS
            , Lerner, L., Costanzo, D., McElligott, K., Boozer, S., Mays, R., Smith
            ,E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
            Creation of genome-wide protein expression libraries using random
   TITLE
            activation of gene expression
            Nat. Biotechnol. 19 (5), 440-445 (2001)
   JURNAL
            21227151
   MEDLINE
            11329013
    PUDMED
            Contact: Scott J. Cain
 COMMINT
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 515.
                     Location/Qualifiers
 FEATURES.
                     1. .784
      source
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="HT1080"
                     /clone_lib="Athersys RAGE Library"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
```

Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT ORIGIN 265 a 151 c 158 g 209 t 1 others

Query Ma Best Loc Matches	cal S	imilarity	66.3%; 96.5%; tive	Pred.	671.8; No. 3.86 smatches	e-135;	Length Indels	784; 3;	Gaps	3;
ДУ		ATGAAAAAAAT ATGAAAAGAAT	1111111	111111			111111	11111		60 91
Qy Db	61 .	CTGAAAGACAA CTGAAAGACAA	TTTGGCC <i>I</i>	ATTTTGG <i>I</i>	AAAGCAA(GACAAAA 	AGACAGACA	AAGGCT	TCAGAA	
Ďρ. ĜÃ		GÄAGTGTCTAA GAAGTGTCTAA			[111111	1 1 1 1 1 1	11111	
Ω χ 200		CCCCCAACAGA CCCCCAACAGA	111111	11 111.	1	! [
98 .		ACACTGATAGO				111111		$ \cdot \cdot \cdot \cdot $	11111	
TOY	12.47	TTTAACAACAT					_	11111	11111	360: 391•
SM T		GCTCATCCTCA	111111	1 1 1 1 1 1		3111111		11111	-111	
Qy Dic		CGTTCTGGGA'		111111						
Qy Do	4 .	TTTTCTAATC	1111111			11111	{			
Qy -		TCAGATGCCT	1111111	111111	11111111		111111		;	
QY	601	TTCTTAGAAC	AAAATTAC	CGACACTA	ATTTTTGA <i>I</i>	AGACTAT	G-AGAAAT1 	rgcttca 	AGTCTGA	659
Db Qу		GAATTATGTT	AC-TAAGA		CTTTAAAGO	TGCTAG		ATCCTG	GACCGTC	

```
692 GAATTATGTTACTTAAGAGACAGTCTTTAGAGCTGCTAGGGGAGCTGATCCTGAAANGTT 751
Db
          719 ACAACTTTGCCATCATGACAAAGT 742
Q۷
               111111111111111111111111111
          752 ACAACTTTGCC-TCATGACAAAGT 774
Db
RESULT 1.0
                                                                   HTC 05-DEC-2002
AK013161
                                                          linear
                                      1281 bp
                                                  mRNA
            Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
LOCUS
             enriched library, clone:2810425013 product:MO25-LIKE PROTEIN
DEFINITION
             homolog [Homo sapiens], full insert sequence.
             AK013151
ACCESSION
             AK013161.1 GI:12850350
VERSION
             HTC; CAP trapper.
KEYWORDS
             Mus musculus (house mouse)
SOURCE
             Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  ORGANISM
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
             Carninci, P. and Hayashizaki, Y.
   AUTHORS
             High-efficiency full-length cDNA cloning
   TITLE
             Meth. Enzymol. 303, 19-44 (1999)
   JOURNAL
             99279253
   MEDLINE
             10349636
    PUBMED
             Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K. K.
 REFERENCE
              Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
   BUTTORS
             Normalization and subtraction of cap-trapper-selected cDNAssto
              prepare full-length cDNA libraries for rapid discovery of new genes
    Genome Res. 10 (10), 1617-1630 (2000)
   CORNAL
              20499374
   MEDLINE
              11042159
   DUBMED
              Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 PETERENCE
              Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
   JUTHORS
              Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
              Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
              Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
              Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
              Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
              RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer
    TITLE
              Genome Res. 10 (11), 1757-1771 (2000)
    JOURNAL
               20530913
    MEDLINE
              11076861
     PUBMED
              Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
  REFERENCE
              Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
    AUTHORS
              Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
               Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
               Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
               Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
               Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
               Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
               Wagner, L., Wasnio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
```

```
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
           Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
           Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
           Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
           Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
           Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
           Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
           Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
            Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
            and Hayashizaki,Y.
            Functional annotation of a full-length mouse cDNA collection
 TITLE
            Nature 409 (6821), 685-690 (2001)
 JOURNAL
            21085660
 MEDLINE
            11217851
  PUBMED
REFERENCE
            The FANTOM Consortium and the RIKEN Genome Exploration Research
  AUTHORS
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
  TITLE
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
  JOURNAL
               (bases 1 to 1281)
REFERENCE
            Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
  AUTHORS
            Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
            Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
            Giraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
            Lasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda. M.,
            Hoya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
            Momura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
            Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
            Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
            (Izuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
            Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
            Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
            Direct Submission
  TTTLE
            Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
  JOSPNAL
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
            Please visit our web site (http://genome.gsc.riken.go.jp/) for
COMMENT
             further details.
             cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues. First strand cDNA was primed with a primer
             [5' GAGAGAGAGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
             and subsequently enriched for full-length by cap-trapper. cDNA went-
             through one round of normalization to Rot = 7.5 and subtraction to
             Rot = 37.5. Second strand cDNA was prepared with the primer adapter
             with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
```

Host: SOLR.

```
Location/Qualifiers
FEATURES
                1. .1281
   source
                /organism="Mus musculus"
                /mol type="mRNA"
                /strain="C57BL/6J"
                /db_xref="FANTOM_DB:2810425013"
                /db_xref="MGI:1908997"
                /db xref="taxon:10090"
                /clone="2810425013"
                /tissue_type="whole body"
                /clone_lib="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="10, 11 days embryo"
                289. .1096
    misc feature
                /note="MO25-LIKE PROTEIN homolog [Homo sapiens]
                (SWISSPROT | Q9H9S4, evidence: FASTY, 98.2%ID, 100%length,
                match=1002)
                putative"
                /db xref="MGI:1922871"
                                 332 t
                   Ź75 ¢
                          291 g
            333 a
BASE COUNT
ORIGIN
                          Score 622.8; DB 11;
                                           Length 1281;
                    61.4%;
  Query Match
                          Pred. No. 1.7e-124;
                    82.8%;
 Best Local Similarity
                                                            2;
                                       97; Indels
                                                  62;
                                                     Gaps
                            Mismatches
                         0:
        763; Conservative
 Natches
         QV
           Db
        212 AATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCTCAGCTAGCACA 212
           AATTCTGTGTGGAACGAACGACAAGGAGCCCCCTACAGAAGCAGTGGCTCAGCTGGCGCA 35690
Dig.
        3 AGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAGCTGATAGACTT 272
                          357 GGAGCTCTACAGCAGCGGGTTGCTGGTGACACTCATAGCTGACCTGCAGCTCATAGACTT 416
Eij.
        273 TGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGACAGATAGGCAC 332
Qŷ
           417 TGAGGCAAAAAAAGATGTGACCCAGATATTCAACAACATCCTGAGAAGACAGATTGGTAC 476
Db
        333 TCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAA 392
            111 HERRICH BERCH 14 HE CHRÜCH MICHERRECH BRIE
 Qy.
        477 ACGGTGTCCTACTGTCGAGTACATCAGTTCTCATCCTCACATCCTGTTTATGCTTCTCAA 536
 Db
        303 AGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCG 452
            Çγ
        537 AGGCTATGAAGCCCCACAGATTGCCTTACGCTGTGGGATTATGCTAAGAGAGTGTATTCG 596
 Db:
        453 ACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTA 512
            97
        597 ACATGAGCCACTTGCCAAAATCATCCTATTTTCTAATCAGTTCAGAGATTTCTTCAAGTA 656
 Db
         513 CGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACT 572
 Qу
             657 TGTTGAGCTGTCCACCTTTGATATCGCTTCAGATGCCTTCGCTACTTTTAAG----- 708
 Db
```

```
573 AACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGA 632
Qу
                                                        -ATTTTTGA 716
Db
        632 AGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCT 692
Qу
            717 AGACTATGAGAAACTGCTGCAATCTGAGAACTATGTGACAAAGAGACAATCTTTAAAGTT 776
Db
        693 GCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAA 752
Qу
            777 GCTAGGTGAGCTGATCCTGGACCGCCACAATTTCACCATTATGACCAAG--TATCAGCAA 834
Db
        753 GCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTT 812-
Qу
            335 GCCAGAGAACCTGAAACTGATGATGAACCTGCTTCGAGACAAAAGTCCCAACATCCAATT 894
Db
        813 TGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGT 872
QУ
             895 CGAAGCCTTCCATGTCTTTAAGGTGTTTGTGGCCAGCCCCCACAAAACGCAGCCTATCGT 954
Db
        873 GGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGA 932
QУ
            955 GGAGATTCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTTCAGAAAGA 1014
Db
        933 AAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGA 992
QУ
            1015 AAGGACAGACGACGAGCAGTTTGCTGACGAGAAGAACTACCTGATTAAACAGATTCGAGA 1074
Do
         993 CTTGAAGAAAACGGCCCCTTGA 1014
            1111111111 | 11111
        1075 CTTGAAGAAAGCAGCCCCGTGA 1096
RESULT 11
BU116523
                                                       EST 25-NOV-2002
                                               linear
                                951 bp mRNA
LOCUS
          603139786F1 CSEQCHL15 Gallus gallus cDNA clone ChEST129122 5', mRNA
DEFINITION
          sequence.
          £U116522
ACCESSION
          BU116522.1 GI:25323402
VERSION
KEYWORDS
          EST.
          Gallus gallus (chicken)
SOURCE
  ORGANISM
          Gallus gallus
          ⊞ukaryota; Metazoa; Çhordata; Craniata; Vertebrata; Euteleostomi;
          Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
           Phasianinae; Gallus.
             (bases 1 to 951)
REFERENCE
          Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
  AUTHORS
           Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
           A Comprehensive Collection of Chicken cDNAs
  TITLE
          Curr. Biol. 12 (22), 1965-1969 (2002)
  JOURNAL
  MEDLINE
           22335534
           12445392
   PURMED
           Contact: Simon Hubbard
COMMENT
           Department of Biomolecular Sciences
           University of Manchester Institute of Science and Technology (UMIST
```

```
PO Box 88, Manchester, M60 1QD, UK
         Tel: 01612008930
         Fax: 01612360409
         Email: Simon. Hubbard@umist.ac.uk.
                 Location/Qualifiers
FEATURES
                 1. .951
   source
                 /organism="Gallus gallus"
                 /mol type="mRNA"
                 /strain="Compton Line 15I"
                 /db xref="taxon:9031"
                 /clone="ChEST129122"
                 /sex="Female"
                 /tissue type="cerebrum"
                 /dev stage="adult"
                 /lab_host="DH10B"
                 /clone_lib="CSEQCHL15"
                 /note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
                 EcoRI; Site_2: Not1; Modification of pBluescript II KS(+)
                 [Stratagene] vector to accommodate cDNA produced with the
                 T-trimmed protocol (Construction of uni-directionally
                 cloned cDNA libraries from messenger RNA for improved 3'
                 end DNA sequencing by Glenn Fu, et al. U.S. Pátent # 6,387
                 ,624). Cut pBluescript II KS(+) with NotI and EccRI.
                 Ligate in double stranded adaptor centaining BsgI and
                 BamHI sites [5'ggccgcgtgcagccccggatccgaaaaaaag]
                 [5'eattctttttttggatccggggctgcacgc] "
                                                        100
                                   240 €
                    305 c
                            202 q
             303 a
BASE LOUIS
ORTGIN
                            Score 614; DB 13; Length 951;
                     60.68;
 Query Match
                                                         - ---
                     32.8%; Pred. No. 1.3e-122;
 Best Local Similarity
                                                        Gaps
                           0: Mismatches 150; Indels
                                                     3:
 Matches //37; Conservative
        \Omega Y
           6 GAAGTGTCAAAATCTCTGCAAGCAATGAAGTAAATTCTGTGTGGGACCACAGACAAGGAG 65
3)55
        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Эy
           66 CCACCGACAGAAGTAGTGGCTCAGCTGGCACAAGAATTGTACAACAGTGGCCTTCTAGTG 125
áC
        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
ŲΨ
            126 ACACTTATTGCCAACCTGCAGCTCATAGATTTTGAGGGTAAAAAGGATGTTTCCCAGATA 185
ĎЬ
        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
ΩŸ
            186 TTTAACAACATCCTGAGAAGACAAATTGGCACACGAAGCCCTACTGTGGAATACATTAGT 245
Dp,
        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
ĈУ
            305 GCCCATCCACATATCCTGTTCATGCTTCTGAAAGGCTATGAATCCCCAAATATTGCCTTA
Db
        421 CGTTGTGGGATTATGC'ÍGAGAGAATGTATTCGACATGAACCACTTGCCAA-AATCATCCT 479
QУ
```

206 CGCTGTGGAATTATGCTGAGGGAGTGCATCCGACATGAACCATTGGCCAACAATCATACT 365

Db

Qу	480 CTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGC 539
Db	366 TTTTTCAGAACAGTTCAGAGACTTCTTCAAGTATGTGGAAATGTCAACATTTGATATAGC 425
Qy	540 TTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGA 599
Db	426 ATCTGATGCCTTTGCTACATTCAAGGACTTGTTAACAAGGCACAAGTTGTTGGTAGCAGA 485
Qy	600 CTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGA 659
Db	486 TTTTATGGAACAAATTATGATACGATCTTTGAGGATTATGAAAAACTCCTTCATTCTGA 545
Qу	GAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCA 719
Db	346 GAATTACGTAACAAAGAGACAGTCTTTGAAGCTGCTGGGTGAATTGATTCTAGACAGAC
Qy	720 CAACTTTGCCATCATGACAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAA 779
Db	606 CAACTTCGCCATCATGACAAAATATATCAGCAAACCAGAGAATCTGAAGCTGATGATGAA 665
Çy	780 CCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTT 839
වුර	656 CTTGCTGCGAGACAAAAGCCCCAACATTCAATTTGAAGCATTCCATGTGTTCAAGGTTTT 725
Qy	©40 PGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 899
ph ·	726 TGTGGCCAGTCCAAACAAAACTCAGCCCATCGTGGAGATCCTGCTGAAAAACCAG-CCAA 784
Egr.	#00 ACTCATTGAGTTTCTGAGCAGCTTCCA-AAAAGAAAGGACGGATGATGAGCAGTTCGCTG 958
Db	785 GCTCATCGAGTTTCTGAGCCATTTCCAGAAACGAGAGGACGGTTGACGAGCAGTTCACCG 844
οχ	959 ACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCC 1008
Db	245 ACGAGAAGAACTACCTGATCAAGCAAATCCGAGACTTGAAGAAGGCCGAC 894
RESULT 12	
DQ669953 LOCUS	BQ669953 982 bp mRNA linear EST 15-JUL-2002
DEFINITION	
ACCESSION	5', mRNA sequence. 3Q669953
VERSION	3Q669953.1 GI:21780787
KEYWORDS	EST.
SOURCE	Homo sapiens (human) 4 Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	Mammalia; Eutheria; Primates; Catalinini; Hominidae; Homo. 1 (bases 1 to 982)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL COMMENT	Unpublished Contact: Robert Strausberg, Ph.D.
· 	Email: cgapbs-r@mail.nih.gov
	Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

```
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCM2407 row: m column: 13
          High quality sequence stop: 508.
                 Location/Qualifiers
FEATURES
                  1. .982
    source
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6255924"
                  /tissue_type="epidermoid carcinoma, cell line"
                  /lab host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_102"
                  /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
                  Site_2: EcoRI; cDNA made by oligo-dT priming.
                  Directionally cloned into EcoRI/XhoI sites using the
                  following 5' adaptor: GGCACGAG(G). Library constructed
                  by Ling Hong in the laboratory of Gerald M. Rubin
                  (University of California, Berkeley) using ZAP-cDNA
                  synthesis kit (Stratagene) and Superscript II RT (Life
                  Technologies). Note: this is a NIH MGC Library."
                                             11 others
                                     197 t
                             357 q
                      200 C
              217 a
BASE COUNT
ORIGIN
                             Score 594.2; DB 13; Length 982;
                      58.6%;
  Query Match
                             Pred. No. 2.6e-118;
                      96.5%;
  Destructal Similarity
                                               Indels
                             0; Mismatches
                                           20;
  Modeles 523; Conservative
         334 CGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAA 393
25.
            1 CGGAGTCCTACTGTGGAG-ATATTAGTGCTCATCCTCATATCCTGGTTATGCTCCTCAAA 59
         394 GGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGA 453
            50 GGATATGAAGCCCCACAGATTGCCTTACATTGGGGGATTATGCTGAGAGAATGGATTCGA 119
מניב
         454 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 513
             QУ
         220 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 179
T^{1}
         514 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 573
             180 GTGCAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 239
 Ĺij
         574 ACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 633
             QУ
         240 ACCAGACATAAAGTGGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 299
 D.b
         534 GACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTG 693
             ***********************************
 QV
         300 GACTATGAGAAATTGCTTCAGTCTGAGAATTATGGTACTAAGAGACAGTCTTTAAAGCTG 359
 Db
         694 CTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAG 753
             Qy.
```

```
360 CTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAG 419
Db
                  754 CCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTT 813
                          QУ
                  420 CCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTT 479
Db
                  814 GAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTG 873
                          QУ
                   480 GAAGCCTTTCATGGTTTTAAGGGGGTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTG 539
Db
                   874 GAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAA 933
                          QУ
                   540 GAGATCCTGGTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAAGAA 599
Db
                   934 AGG--ACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAAC 982
                                       THEFT IS THE TOTAL TOTAL
 QY
                           111
                   600 AGGGACGGGATGATGAGCANNTCCCTGACGAGAAAGACTACTTGGGTTAAC 650
 Db
 MESULT 13
 3U518807
                                                                                                                      EST 12-SEP-2002
                                                                                                      linear
                                                                                       mRNA
                                                                     934 bp
                       BU518807
 LOCUS
                      AGENCOURT_10171930 NIH_MGC_134 Mus musculus cDNA clone
 DEFINITION
                       IMAGE:6516567 5', mRNA sequence.
                       BU518807
 ACCESSION
                       BU518807.1 GI:22826333
 TESION
                                                                                                                            KETTIN
                       ST.
    TWORDS
                                                                                                                            BOURTE
                       Fus musculus (house mouse)
 HOUSE CE
                                                                                                                               0.9(1)
                      musculus
     ANISM
                       mukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                        (bases 1 to 934)
  LEWINER NCE
                       NIH-MGC http://mgc.nci.nih.gov/.
     AUTHORS
                       Mational Institutes of Health, Mammalian Gene Collection (MGC)
     TITLE
                        Unpublished
     COURNAL
                                                                                                                              17.15W} ......
                        Contact: Robert Strausberg, Ph.D.
 COMMENT
                        Email: cgapbs-r@mail.nih.gov
                        Tissue Procurement: Dr. David Rowe
                          cDNA Library Preparation: Invitrogen Corp
                          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                          DNA Sequencing by: Agencourt Bioscience Corporation
                         Clone distribution: MGC clone distribution information can be
                        found through the I.M.A.G.E. Consortium/LLNL at:
                        http://image.llnl.gov
                        Plate: LLAM14095 row: e column: 16
                        High quality sequence stop: 656.
                                         Location/Qualifiers
   FEATURES
                                         1. .934
            source
                                         /organism="Mus musculus"
                                         /mol_type="mRNA"
                                         /db xref="taxon:10090"
                                         /clone="IMAGE:6516567"
                                         /tissue_type="undifferentiated limb"
                                         /lab host="DH10B (phage-resistant)"
                                          /clone lib="NIH_MGC 134"
                                          /note="Vector: pCMV-spORT6.1.ccdb; Site_1: EcoRV; Site_2:
                                         NotI; Cloned unidirectionally. Primer: Oligo dT. Average
```

insert size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library." 301 a 198 c 200 g 234 t 1 others

BASE COUNT

	Match ocal S s 654	imilarity	89.0%;	Pred.	585.2; No. 2.3e matches	:-116;	Length	934;	Gaps	2;
Qy Db		TGAAAAAAATG TGAAAAAAATG	111111	111111	11111	11111	1111111		! } .	
Qy Db		TGAAAGACAAT TGAAAGACAAC		111111	1111111		111111			
Oy Db	122	λ አ ሬጥረጥረጥል Δ Δ	TCACTGC	AGCAATO	GAAAGAAAT	TCTGTG	TGGTACAA	ACGAGA.	AAGAAC	181
Qy	182	CCCCAACAGAA	GCAGTGG(CTCAGCT <i>I</i>	AGCACAAGA	AACTCTA	CAGCAGTG(GCTGC	TAGTGA	241
Db Cy	242	CACTGATAGCT	GACCTGC	AGCTGAT <i>l</i>	AGACTT'IGA	AGGGAAA	AAAAGATG 	TGACCO	AGA TAT	301
ýy Saa	2 A.	TTAACAACATO	CTTGAGAA(GACAGATA	AGGCACTC	GGAGTCC	TACTGTGG	AGTATA	TTAGTG	3610?
1/3.4 	(1) 表表	OTATION THE COMPANY	ATCCTGT	TTATGCT(CCTCAAAG	GATATGA	AGCCCCAC	AGATTC	CCTTAC	121 2
Qy	£22	GTTGTGGGATT	ratgetga	GAGAATG'	TATTCGAC	ATGAACO	CACTTGCCA	AAATCA	ATCCTCT	481
Db C	482	TTTCTAATCAATCTAATCTAATCTAATCAA	ATTCAGAG	ATTTCTT	TAAGTACG	TGGAGT	rgtcaacat	TTGATA	ATTGCTT	541
Qy	542	CAGATGCCTT'	rgctactt	TCAAGGA	TTTACTAA	CCAGAC	ATAAAGTGI	TGGTAC	GCAGACT	601
Db Ωγ	ຍິນ3	TCTTAGAACA	 AAATTACG	ACACTAT	'TTTTGAAC	ACTATG	AGAAATTGO	CTTCAG	rctgaga 	661
Qy	662	ATTATGTTAC ACTATGTGAC	TAAGAG- <i>A</i>	CAGTCTI	TAAAGCTO	CTAGGG	-GAGCTGAT	CCTGG	ACCGTCA	719
Dδ	1.13	ACIAIGIGAC	LATOCOCK							

```
720 CAACTTTGCCA'ICAT 734
Qy.
                 833 CACAATTTTCACCAT 847
Db
```

RESULT 14 CD354831 linear EST 29-MAY-2003 713 bp mRNA CD354831 LOCUS. UI-M-GMO-cge-i-10-0-UI.rl NIH_BMAP GMO Mus musculus cDNA clone DEFINITION IMAGE: 30361641 5', mRNA sequence. CD354831 ACCESSION CD354831.1 GI:31147332 VERSION KEYWORDS EST. Mus musculus (house mouse) SOURCE Mus musculus ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 713) REFERENCE NIH-MGC http://mgc.nci.nih.gov/. AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC) TITLE Unpublished JOURNAL Contact: Robert Strausberg, Ph.D. COMMENT Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at

cDNA Library preparation: Dr. M. Bento Soares, University of Towa http://genome.uiowa.edu/distribution/mousefl.html This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seg primer: pYX-5.

10.5

worm.ida

Location/Qualifiers

1. .713 /organism="Mus musculus" /mol type="mRNA"

/strain="C57BL/6" /db xref="taxon:10090"

/clone="IMAGE: 30361641" /tissue type="whole brain"

/dev_stage="1, 5 and 15 days newborn" /lab_host="DH10B (T1 phage resistant)"

/clone lib="NIH BMAP GMO"

/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site 2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site.Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

BASE COUNT 220 a 159 c 140 g 192 t 2 others ORIGIN

57.1%; Score 579.2; DB 14; Length 713; Query Match Pred. No. 4.4e-115; 89.9%; Best Local Similarity Indels Gaps 2: Mismatches 70: Matches 642; Conservative 0; 282 AAAAGATGTGACCCAGATATTTAACAACAT-CTTGAGAAGACAGATAGGCACTCGGAGTC 340 Qу 1 AAAAGATGTGACCCAGATATTCAACAACATCCNTGAGAAGACAGATTGGTACACGGTGTC 60 Db' 341 CTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATG 400 QY51 CTACTGTCGAGTACATCAGTTCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATG 120 Db 401 AAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAAC 460 ÇУ 121 AAGCCCCACAGATTGCCTTACGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGC 180 Db 461 CACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGT 520 ŐУ 131 CACTTGCCAAAATCATCCTATTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGC 240 Dp 1011 TGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGAC 580 27 2 1 TGTCCACCTTTGATATCGCTTCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGAC 3004 31 ATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATG 640 ATAAAGTATTGGTAGCAGACTTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATG 360 AGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGG 700) Çy AGAAACTGCTGCAATCTGAGAACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTG 420 Db 701 AGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGA 760 Qу 421 AGCTGATCCTGGACCGCCACAATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGA 480 D5 761 ACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCT 820 Qу 481 ACCTGAAACTGATGATGAACCTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCT 540 Db 821 TTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCC 880 Qу 541 TCCATGTCTTTAAGGTGTTTGTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTC 500 Db. Qу Db 941 ATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACT 994 ΩУ 561 ACGACGAGCAGTTTGCTGACGAGAAG-ACTACCTGATTANACAGATTCGAGACT 713 Db

```
RESULT 15
HSM073180
     HSM073180 standard; RNA; EST; 742 BP.
XX
AC
     BX483012;
XX
     BX483012.1
SV
XX
     09-MAY-2003 (Rel. 75, Created)
DT
     09-MAY-2003 (Rel. 75, Last updated, Version 1)
DT
XX
     Homo sapiens mRNA; EST DKFZp686C08234 r1 (from clone DKFZp686C08234)
DΕ
XX
     EST; expressed sequence tag.
KW
XX
     Homo sapiens (human)
OS
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC
     Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC
XX
KN
RP
     1.742
     Ottenwaelder B., Obermaier B., Deutschenbaur S., Mewes H.W., Weil B.,
RA
     Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RA
RT
     Submitted (07-MAY-2003) to the EMBL/GenBank/DDBJ databases.
RL
     MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
ŔЬ
XX.
CH.
     This is the 5' sequence of the clone insert
     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC
    Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC -
     sequenced by MediGenomix (Martinsried/Germany) within the cDNA
CC
     sequencing consortium of the German Genome Project.
CC.
\{ \gamma_{i,j}^{m} \}_{i=1}^{m} .
     Mo s1 sequence available.
     This clone (DKFZp686C08234) is available at the RZPD in Berlin.
CC:
CC
     Please contact the RZPD: Ressourcenzentrum, Heubnerweg 5,
CC
     14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
FH ·
     Кеу
                     Location/Qualifiers
ΞH
E_{\rm eff}
                     1. .742
     source
                     /db xref="taxon:9606"
FT
                     /mol type="mRNA"
FT
                     /organism="Homo sapiens"
                     /clone="DKFZp686C08234"
BT
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
P^{i}T
                     DH10B; sites SfiIA + SfiIB"
7''7'
                     /dev stage="adult"
FT
                     /tissue_type="cDNA-collection"
FŢ.
\mathbf{X}'\mathbf{X}
     Sequence 742 BP; 256 A; 143 C; 162 G; 179 T; 2 other;
SQ
                          57.0%; Score 578; DB 2; Length 742;
  Query Match
                          99.8%; Pred. No. 8e-115;
  Best Local Similarity
  Matches 578; Conservative
                                 0; Mismatches
                                                  1;
                                                        Indels
                                                                              0;
            1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Qу
```

Db	164 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 223
Qy	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC
Db	
Qy	121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
da	
Öλ	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Db	
Ой	2#1 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
DD	TA GAGAMA GEGACTCCGAGTCCTACTGTGGAGTATATTAGT 360
Ç <u>v</u> Db	301 STTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGTGTG
. ,	361 SCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
ζy	361 SCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGMGGGGCCCACAGATTGCCTTA 583
1 * 1 *-11	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
300.4	
. Q;:	681 'TTTCTAATCAATTCAGAGAT'ITCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
Mit.	
	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGA 579
Dþ.	/U4 ICAGAIGCGIIIGGIAGI

Search completed: January 5, 2004, 03:18:09 Cob time : 2589 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model January 6, 2004, 00:37:47; Search time 3965 Seconds Run on: (without alignments) 10462.134 Million cell updates/sec US-10-088-872-1 Title: Perfect score: 1014 Sequence: 1 atgaaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014 Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0 2888711 seqs, 20454813386 residues Searched: Total number of hits satisfying chosen parameters: 5777422 Minimum DB seq length: 0 Maximum DB seq length: 2000000000 et i j Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries GenEmbl:* โดยเปลือง เ 1: gb ba:* 2: gb_htg:* 3: gb_in:* 4: gb_om:* 5: gb_ov:* 6: qb pat:* 7: gb ph:* 8: gb_pl:* 9: gb_pr:* 10: gb_ro:* 11: gb_sts.* 12: gb_sy:* 13: gb_un:* 14: gb_vi:* 15: em_ba:* 16: em_fun:* 17: em_hum:* 18: em_in:*

em_mu:*

em om:*

21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

19:

20:

```
28:
    em_un:*
    em_vi:*
29:
    em_htg_hum:*
30:
     em_htg_inv:*
31:
     em htg_other:*
32:
     em_htg_mus:*
33:
     em_htg_pln:*
34:
     em_htg_rod:*
35:
     em htg_mam:*
36:
     em_htg_vrt:*
37:
38:
     em_sy:*
     em_htgo_hum:*
39:
     em_htgo_mus:*
40:
     em_htgo_other:*
41:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Re	sult _i . Kos	Score	Query Match I	Length	ĎB	ID	·	Description
-	1	1014	100.0	1014	6	AX105381		AX105381 Sequence AR097361 Sequence
	2	1010.8	99.7	1344	5	AR097361		AR203365 Sequence
	3	1010.8	99.7	1344	.6	AR203365		BC010993 Homo sapi.
	4	1009.2	99.5	1491	9.	BC010993		BD157871 Primer fo
	5	992.8	97.9	2002	ŧ,	BD157871		AK022639 Homo sapi
	6	992.8	97.9	2002	9	AK022639		BC016128 Mus muscu
	7	360.4		1359	10	BC016128		BC016546 Mus muscu
	-8	858.8	84.7	1530	10	BC016546		BD147463 Primer to
	3	770.6	76.0	822	. 5	BD147463		BD079551 Cancer-a's
	1.0	004.5	67.5	831	6	BD079551		AX061831 Sequence
	11	32.5	57.5	1026	6	AX061831		AX082322 Sequence
	12	582.6	57.5	3281	6	AX082322		BC020570 Homo sapi
	. 13	582.6	57 .5	3761	9	BC020570		AF151824 Homo sapi
	14	581.6	57.4	1680	9	AF151824		AF113536 Homo sapi
	15	581.6	57.4	3466	9	AF113536		BC020041 Mus muscu
	16	568.8	56.1	1947	. 10			S51858 MO25 gene [
	17	567.2	55.9	2322	10			AK000804 Homo sapi
	18	552.8	54.5	3346	9	AK000804		AX105727 Sequence
	19	541.6	53.4	1053	6	AX105727		AY211923 Homo sapi
	20	533.6	52.6	851	9	AY211923	*	BD079552 Cancer-as
	c 21	520.2	51.3	833	б	BD079552		BC044172 Danio rer
	.2.2.	478.8	47.2	2366	5	BC044172		BC029053 Mus muscu
	23	3,5,8.8	36.4	2991	10			AC101941 Mus muscu
	24	365	36.0	158599		AC101941		AY128438 Drosophil
	25	362.8	35.8	1894	3	AY128438		AC019926 Drosophil
	26	362.8	35.8	8324	2	AC019926		AC010688 Drosophil
	c 27	362.8	35.8	201313		AC010688		AE003526 Drosophil
	c 28	362.8	35.8	281101		AE003526		AB000402 Drosophil
	29	361.2		2213		AB000402		AK026335 Homo sapi
	30	270		1452		AK026335		AJ238632 Chlorella
	31	255		1466		CPR238632		BC034159 Mus muscu
	32	240.8		3514				AY088359 Arabidops
	33	210.4	20.7	1474	. 8	AY088359		11230332

34 35 36 37 c 38 c 39 c 40 c 41 c 42 c 43 44	209.6 200.2 200.2 195 195 186.4 185.4 185.4 185.4 172.8 169.8	20.7 19.7 19.7 19.2 19.2 18.4 18.3 18.3 17.0 16.7		8 8 8 6 6 8 9 6 6 9 6 6	AF428296 AY055792 AF380659 AX333971 AX407834 AC006284 AL138875 AX711879 AX739961 AL136218 AX283846 AX283847	AF428296 Arabidops AY055792 Arabidops AF380659 Arabidops AX333971 Sequence AX407834 Sequence AC006284 Arabidops AL138875 Human DNA AX711879 Sequence AX739961 Sequence AL136218 Human DNA AX283846 Sequence AX283847 Sequence
--	---	--	--	-------------------------	--	---

ALIGNMENTS

```
RESULT 1
                                                                 PAT 30-APR-2001
AX105381
                                                        linear
                                                DNA
                                    1014 bp
            AX105381
LCCUS
            Sequence 1 from Patent W00123552.
DEFINITION
            AX105381
ACCESSION
            AX105381.1 GI:13921508
VERSION
KEYWORDS
            Homo sapiens (human)
SOURCE
            Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  ORCANISM . Homo sapiens
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
PERHRENCE
            den Daas, I. and Duecker, K.
            Human paralogue of a head trauma induced cytoplasmatic dalcium
  AUTHORS
  TE
            binding protein
            Patent: WO 0123552-A 1 05-APR-2001;
  COURNAGE
             MERCK PATENT GmbH (DE)
                     Location/Qualifiers
 MEANURES.
                      1. .1014
      source
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      1. .1014
      CDS
                      /note="unnamed protein product"
                      /codon_start=1
                      /protein_id="CAC37735.1"
                      /db_xref="GI:13921509"
                      /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                      SLQAMKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                      NILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                       FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                       SENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                       FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
                       TAP"
                                             260 t
                                    209 g
                           205 C
                  340 a
 BASE COUNT
```

Query Match 100.0%; Score 1014; DB 6; Length 1014;
Best Local Similarity 100.0%; Pred. No. 6.9e-238;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CRICIN

Qy	1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Db	1 ATGAAAAAATGCCTTTGTTTAGTAAATCTOOTS
Qy	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC
Db	TARANA TERRETETETETETETETA A A CGAGAAAGAA 180
QУ	121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
Db	121 GAAGTGTCTAAATCACIGCAAGCAATGTITTCT
Qy	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Db	TO THE TRACE OF THE TRACE AND A A AGATGTGACCCAGATA 300
Θλ	241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAATTTTTTTT
Db	TO GO GATTA GOGA CTCGGAGTCTACTGTGGAGTATATTAGT 360:
Qγ	361 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCGTTAGT 360 362 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Db	ACCOUNTATION ACCOU
Qy Db	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420:
	TO A THE STATE OF THE COACATE ACCOUNTING CAAAATCATCCTC 480
97 - 198 :	
$\mathcal{Q}_{\mathcal{V}}$	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
વર્ષેદ્ર	431 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTCGAGTTTT
ðλ	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
Db	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCTOTOTOTOTOTOTOTOTOTOTOTOTOTOT
CA	TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
Db	601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGATGATGCTCCACCGTCAC 720
Çy	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
ďŒ	661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGTT
Ōλ	721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
пр	
QY	781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Db	781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTGATG

```
841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
         Qу
      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Db
      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
         Qу
       901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGCACGGATGATGAGCAGTTCGCTGAC 960
Db
       961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
         Qy
       961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Db
RESULT 2
                                            PAT 14-FEB-2001
AR097361
                                DNA
                                      linear
                         1344 bp
        AR097361
LOCUS
        Sequence 2 from patent US 6071721.
DEFINITION
        AR097361
ACCESSION
                GI:12806091
        AR097361.1
VERS10N
KEYWORDS
        'Jnknown.
SOURCE
  ORGANISM
        Unknown.
         Unclassified.
          (bases 1 to 1344)
        Tang, Y. Tom., Guegler, K.J., Corley, N.C. and Gorgone, G.A.
REFERENCE
  AUTHORS
         Calcium binding protein
  TITLE
         Patent: US 6071721-A 2 06-JUN-2000;
  COURNAL
              Location/Qualifiers
 FERRIRES
               1. .1344
    source
               organism="unknown"
                              353 t
                        280 AT
                  261 €
 BASE COUNT
            450 a
 ORIGIN
                                       Length 1344;
                        Score 1010.8; DB 6;
                   99.78;
  Ouerv Match
                        Pred. No. 4.2e-237;
                  99.8%;
  Best Local Similarity
                                                Gaps.
                                      Indels
                       0; Mismatches
                                    2;
  Matches 1012; Conservative
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          QY
        124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 183
 Db
        Q٧
        of
        QУ
        Db
        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           QV
        304 CCCCCGACAGAACCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363
 Db
        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           Qy
        354 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423
 Db
```

教養衛

Qy	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Db	424 TTTAACAACATCTTGAGAAGACACATTTAG
Qу	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Db	THE CANCEL AND CONCERN CONCERN AND CONCERN CONCERN AND
Ćλ	421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGGGTATTGGTATTCGACATGAACCACTTGGGTATTGGTTATTCGACATGAACCACTTGCCAAAATCATCCTC 603. 544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603.
Db	THE ACTION COTTON OF THE TOTAL ACAPTITISATATICCT 540
QУ	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGIGGAGIIGICIAIGIIIIIIIIII
Db	COLD A COLD A A GTGTTGGTAGCAGAC 600
QУ	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGCTAGCAGACATAAAGTGTTGCTAGCAGACATAAAGTGTTGGTAGCAGAC 723 664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723
Db	CASTATCACA AATTGCTTCAGTCTGAG 660
QУ	TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGGTTGAGAAATTGGTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGGTTCAGTCTGAG 783.
Db	TOTAL TRANSPORT TO THE
ÖÄ.	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGTAGGGGAGCTGATCCTGGACCGTCAC 843.3
7b ·	TRACCARACTCATGATCATCATCATCATCATCATCATCATCATCATCATCATCA
ÖΆ	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTTCATGATGAACCGGAGAACCTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGATGAACCGGAGAACCTGAAACTCATGAACCAAACCAAACCAAACAAA
Db	TO CHIMPON ACCOUNT CATCATTATAAGGTGTTT 840
QΣ	731 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTAAGGTGTTT 963 204 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963
ĎĎ	841 GTGGCCAGTCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
ØΣ ≥ −	841 GTGGCCAGTGCTCTATAAAAACCCCAAA 1023
Db	901 CTCATTGAGTTCTGAGCAGCTTCCAAAAAGAAAGACGGATGATGAGCAGTTCGCTGAC 960
QУ	901 CTCATTGAGTTTCTGAGGAGTTCTGAGCAGCTTCCCAAAAAGAAAG
ďŒ	961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QΥ	1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
Dр	1084 GAGAACAACIIICIII
AR20 LOCU DEFI ACCI	INITION Sequence 2 from patent ob osser ESSION AR203365
KEY	WORDS
SOU	RCE Unknown.

```
Unknown.
 ORGANISM
        Unclassified.
          (bases 1 to 1344)
        Tang, Y. Tom., Guegler, K.J., Corley, N.C. and Gorgone, G.A.
REFERENCE
 AUTHORS
        Calcium binding protein
        Patent: US 6365371-A 2 02-APR-2002;
 TITLE
 JOURNAL
              Location/Qualifiers
FEATURES
              1. .1344
   source
              /organism="unknown"
                             353 t
                 261 C
                       280° g
           450 a
BASE COUNT
ORIGIN
                                      Length 1344;
                                 DB 6;
                       Score 1010.8;
                  99.7%;
 Cuery Match
                       Pred. No. 4.2e-237;
                  99.8%;
 Best Local Similarity
                                                    0;
                                               Gaps
                                            0;
                                     Indels
                                   2:
                         Mismatches
                       0;
           Conservative
 Matches 1012;
        1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          QУ
      124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
Db
        QУ
       Db
       181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
          \bigcirc \lor
        304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363
 1
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
          CV.
        364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423
 Db
        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           Ωy
        424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483
 DD.
        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           QΥ
        484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543
 Db
        421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           Çy
          CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603
 Tib
        481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540....
           QУ
        504 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663
  Dr.
        541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
           QУ
         564 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC
  Db
         601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
  QУ
```

7)		 CTGAG 783
Db	THE CONTROL OF THE CO	CGTCAC 720
QУ	661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACG	CGTCAC 843
Db	784 AATTAIGTTACTAAGAGTGATTG	አጥሮአአሮ 780
Qу	721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATG.	ATGAAC 700. ATGAAC 903 .
ďQ	844 AACTTTGCCATCATGACAAAGTATATCAGGTTGGAACCCTTTCATGTTTTTAAG	GTGTTT 840
QУ	781 CICCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAG	 GTGTTT 963
Db	204 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTT	aggann 900
Õλ	343. GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAG 	 CCCAAA 1023
Db	964 GTGGCCAGTCCTCACAAAACACAGGGTT	
QУ	901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTC	GCTGAC 960 CGCTGAC 1083
Dp	1024 CTCATTGAGTTTCTGAGCAGCTTCCAURE	- ::014
Ç.A	961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTG. 	A 1014
DD:		
DESCUR DESCUS	1491 bp mRNA linear PRI SC010993 typethotical protein FLJ12577, clone MGC:	25-JUL-2001 15031
DEFINITI	ION Homo sapiens, hypothetical public interest in IMAGE: 3956127, mRNA, complete cds.	
ACCESSIO	on BC010993	*
VERSION	~ ~ ~ 1 CO10177	and the second of the second o
KUYWORDS	g MGC.	in the second
SOURCE	Homo sapiens (human)	A CONTRACTOR
ORGANII	IISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu	teleostomi;
	Eukaryota; Metazoa; Chordata; Claniata, Versosania, Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homin	omo.
	. 1 La 3/011	•
REFERENC AUTHOR		
TTTLE	Direct Submission Notice of Health,	Mammalian
JOURN	1 (02 TIT ONGT) NACTORAL INSCIDENCES	l Cancer 0892-2590,
	USA NIH-MGC Project URL: http://mgc.nci.nih.gov	A STATE OF S
REMAR	RX NIH-MGC Project URL: http://mgc.mgz.mg	#1W
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov	
	. n	
		.MI.)
		17.A.T.)
	DWA Commencing by: Institute for Systems 22	
	http://www.systemsbiology.org	
	contact: amadan@systemsbiology.org Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie	Ford, Julia

Greene, Mark Ketteman and Anuradha Madan

```
Clone distribution: MGC clone distribution information can be found
        through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
        Series: IRAL Plate: 25 Row: k Column: 12
        This clone was selected for full length sequencing because it
        passed the following selection criteria: matched mRNA gi: 10434146.
               Location/Qualifiers
FEATURES.
               1. .1491
               /organism="Homo sapiens"
   source
               /mol type="mRNA"
               /db_xref="LocusID:81617"
               /db_xref="taxon:9606"
               /clone="MGC:15031 IMAGE:3956127"
               /tissue_type="Placenta, choriocarcinoma"
               /clone_lib="NIH_MGC_21"
               /lab_host="DH10B-R"
               /note="Vector: pOTB7"
               416. .1285
    CDS
               /codon start=1
               /product="hypothetical protein FLJ12577"
                /protein_id="AAH10993.1"
                /db_xref="GI:15012173"
                -translation="MKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFE
                EKKDVTQIFNNILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECI
                RHEPLAKIILFSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT
                IFEDYEKLLQSENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKS
                PNIQFEAFHVFKVFVASPHKTQPIVETLLKNQPKLIEFLSSFQKERTDDEQFADEKNY
                LIKOIRDLKKTAP"
                                                  Angle of David
                                393 t
                   290 c
                          305 g
                                                   1 35 3
            503 a
 MOGE COUNT
 ORIGIN
                                                   , 4€
                                          Length 1491:
                          Score 1009.2; DB 9;
                    99.5%;
  Caracy Mateh
                         Pred. No. 1e-236;
                    99.7%;
  Best Local Similarity
                                                    Gaps :-
                                         Indels
                            Mismatches
                                      3:
  Motones 1011; Conservative
                         0.;
          ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
           270 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 331
 QΥ
         Db
            QУ
 Do
         QУ
         Dio
         181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240 ~
            QУ
         452 CCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 511
         241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
  Db
            QУ
         301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
  Db
```

QУ

Db	
Qу	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420 1
Db	632 GCTCATCCTCATATCCTGTTTATOGTO
Qγ	421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480 421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 751
Db	THE CONTRACT CAPTURE OF THE CONTRACT CAPTURE OF THE
Ο̈́А	
Db	TO GA GA GA A A CTG TTGGTAGCAGAC 600
Ğλ	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGTTGTTGCTACTTCAAGGATTTACTAACCAGACATAAAGTGTTGTTGCTAGCAGAC 871 812 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 871
מֹת	THE THE HEALT CAN A THE COURT CAG TO COME TO THE CAG TO
·Çy	
Dþ	THE GOOD COTTON TO THE TOTAL OF
QY	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATGGTGGTGGTGGTGGTAGGGGAGCTGATGGTTAGAGCTGCTAGGGGGAGCTGATCCTGGACCGTCAC 991 2 5 922 PATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 991 2 5 922 PATTATGTTACTAAGAGACAGTCTTTAAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 991 2 5 922 PATTATGTTACTAAGAGACAGTCTTTAAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 991 2 5 92 92 92 92 92 92 92 92 92 92 92 92 92
. D'5 -	A DECEMBER OF A COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND COMMAND AND COMMAND AND COMMAND AND COMMAND AND COMMAND AND CO
	A AGGRETTETAAGGTGTT 040:: .
\mathcal{D}_{λ}	
Do	THE CAMPACTOTTA A A A A A A TCAGCCCAAA 900
्रेट्र	841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAATCAGCCCAAA 1171 GTGGCCAGTCCTCACAAAACACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAA 1171 GTGGCCAGTCCTCACAAAAACACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAA 1171
Db	1112 GTGGCCAGTCCTCHALL 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Qy	901 CTCATTGAGTTTCTGAGCAGCTTCCAAAATTOTTTTTTTTTT
ΟÞ	961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Ö.Ā.	963 GAGAAGAACTACTTGATTAAACAGATCCGAGAGTTIIIIIIIIII
ប់២	1232 GAGAAGAACTACTIO
	ULT 5 57871 2002 bp DNA linear PAT 17-JAN-2003
. عد بروبيد	2002 bp Division use thereof.

```
LCCUS D157871 Z00Z DD DNA THE TOTAL STREET TO SYNTHESIZING FULL-length cDNA and use thereof.

ACCESSION ED157871
VERSION BD157871.1 GI:27863629
KEYWORDS JF 2002191363-A/12714.
SOURCE Homo sapiens (human)
ORCANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
REFERENCE
          Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 AUTHORS
          Primer for synthesizing full-length cDNA and use thereof
          Patent: JP 2002191363-A 12714 09-JUL-2002;
 TITLE
 JOURNAL
          HELIX RESEARCH INSTITUTE
               Homo sapiens (human)
          OS
COMMENT
               JP 2002191363-A/12714
           PN
               09-JUL-2002
           PD
               28-JUL-2000 JP 2000280990
               TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
           PF
           PΙ
               JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
           PΙ
           PΙ
               KEIICHI NAGAI, TETSUJI OTSUKI
           PΙ
           C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
               C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
            10,
            Primer for synthesizing full-length cDNA and use thereof FH
           PC
                     Location/Qualifiers
                              (127) \cdot (993).
                CDS
            TT
                   Location/Qualifiers
 FEATURES
                   1. .2002
                    /organism="Homo sapiens"
     BOUTGE
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                                       527
                               463 g
                       418 C
               594 a
 BASE COUNT
 ORIGIN -
                                                  Length 2002;
                               Score 992.8; DB 6;
                        97.98;
                               Pred. No. 1.1e-232;
   Chary Match
                        99.88;
                                                                     0;
                                                              Gaps
   Bost Local Similarity
                                                  Indels
                                  Mismatches
                               0;
           994; Conservative
   Matches
           19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
              1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
  C_{I}
  Do
            79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
              61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120
  QУ
           139 CAAGCAA'GAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 198
  Db
               121 CAAGCAATGAAAGAAATTCTGTGTGTGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 180
  QΨ
           199 GCTCAGCTAGCACAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
   Do
               181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240
   QУ
           259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
   Db
               241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300
   QΥ
            319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
   Db.
                QУ
```

	301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360
Db	301 AGACAGATAGGCACTCGGAGTCCTACTGTGGGATTATGCTGTGGGATTATGCTG 438
Qу	379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
Db	261 TTTATGCTCCTCAAAGGATATGAAGCCCCTGTTTTT
QУ	439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498 439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCATCTCTTTCTAATCAATTCAGA 480. 421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480.
Db	421 ACAGAATGTATTCGACATGAACCACTTOTCM2222
QУ	499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
da	A CONCURRENCE TARGET TO THE STATE OF THE STA
Qy 	559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
Db .	THE STATE OF THE S
Qy	619 GACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTTTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTTTTTTT
Do	THE TRANSPORT OF THE TR
QY	
DB	TO THE TRANSPORT OF THE PROPERTY OF THE PROPER
Qy Db	
e de la companya de l	A COMPACA TOTT TTTA AGGTGTTTTTGTGGCCAGTCCTCACAAA 858
Qy ∋b	
ŌΥ	ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 918
ďŒ	OAL ACACACCTATTGTGGAGATCCTGTTATT
Ωу	919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
Db	
Qy Db	979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
, D.	
RESULT AK0226 LOCUS DEFIN	AK022639 2002 bp mRNA linear PRI 01-AUG-2002 AK022639 TOWN FLIT12577 fis, clone NT2RM4001047, highly similar
ACCESS VERSION KEYWO SOURC	SION AK022639 ON AK022639.1 GI:10434146 RDS cligo capping; fis (full insert sequence). E yomo sapiens (human)
ORG	ANISM Homo sapiens

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
REFERENCE
            Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,
  AUTHORS
            Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
            Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,
            Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,
            Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,
            Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
            NEDO human cDNA sequencing project
  TITLE
            Unpublished
  JOURNAL
            2 (bases 1 to 2002)
REFERENCE
            Isogai, T. and Ctsuki, T.
  AUTHORS
             Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
  TITLE
             Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
   JOURNAL
             (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
             NEDO human cDNA sequencing project supported by Ministry of
             International Trade and Industry of Japan; cDNA full insert
 COMMENT
             sequencing: Research Association for Biotechnology; cDNA library
             construction, 5'- & 3'-end one pass sequencing and clone selection:
             Helix Research Institute (supported by Japan Key Technology Center
             etc.) and Department of Virology, Institute of Medical Science,
             University of Tokyo.
                      Location/Qualifiers
  STATURES
                       1. .2002
                       /organism="Homo sapiens"
      SOURC.
                       /mol type="mRNA"
                       /db_xref="taxon:9606"
                       /clone="NT2RM4001047"
                       /cell_line="NT2"
                       /cell_type="teratocarcinoma"
                       /clone_lib="NT2RM4"
                       /note="cloning vector: pME18SFL3~mRNA from uninduced NT2
                       neuronal precursor cells."
                       127. .996
                       /note="unnamed protein product"
       CDS
                        /codon_start=1
                        /protein_id="BAB14147.1"
                        /db_xref="GI:10434147"
                        /translation="MKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFE
                        GKKDVTQIFNNILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECI
                        RHEPLVKIILFSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT
                        IFEDYEKLLQSENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKS
                        PNIQFEAFHVFKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNY
                        LIKQIRDLKKTAP"
                                              527 t
                            418 C
                                     463 g
                   594 a
   BASE COUNT
   ORIGIN
                                                          Length 2002;
                                     Score 992.8; DB 9;
                             97.98;
     Query Match
                                      Pred. No. 1.1e-232;
                             99.8%;
     Best Local Similarity
                                                                         Gaps
                                                           Indels
                                        Mismatches
     Matches 294; Conservative
                                     0;
              19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
                  QУ
                1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
    Db
```

	TAGA GA AGGCTTCAGAAGAAGTGTCTAAATCACTG 138
ΔÄ	79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
Dp	61 ATTTTGGAAAAGCAAGTG 198
QΥ	61 ATTTTGGAAAGCAAGACAAATTOTTCTGTGTGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 198 139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 180 121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
Db	121 CAAGCAATGAAAGAAATTCTGTGTGTGTGACACACTGACACACTGACACACTGACACACTGACACACTGACACACTGACACACAC
QŸ	121 CAAGCAATGAAAGAAATTCTGTGTGTGTGTGACACTGATAGCTGACCTG 258 199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258 191 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240 181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACCTCTTGAGA 318
Dp	181 GCTCAGCTAGCACAAGTETE
QY	181 GCTCAGCTAGCACAAGAACTOTTOM 259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318 259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300 241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 378
Do	241 CAGCTGATAGACTITGAGGGCTCATATCCTG 378
Οğ	241 CAGCTGATAGACTTTGAGGGATTTTTTTTTTTTTTTTTT
' pþ.	
Qy	301 AGACAGATAGGCACTCGGAGTCCTTTACTTACTTGTGGGATTATGCTG 438 379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
dg.	361 TTTATGCTCCTCAAAGGTTCAGA 498
ŨΣ	361 TTTATGCTCCTCAAAGGATATGATCATCATCAAAATCATCCTCTTTTCTAATCAATTCAGA 498 339 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 480 421 AGAGAATGTATTCGACATGAACCACTTGTCAAAAATCATCCTCTTTTCTAATCAATTCAGA 480
	421 AGAGAATGTATTCGTACT 558
Qy	421 AGAGAATGTATTCGACATGATASSA 493 CATTTCTTTAAGTACGIGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558 493 CATTTCTTTAAGTACGIGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540 481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACAACAAATTAC 618
$\mathfrak{D}_{\frac{1}{2}}$	GATTICITIANO DE LA CITATRICITA GARACANA ATTAC 618
Qy	TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600
D, 3	
Ω¥	TTCAAGGATTTACTAACCAGACATTACTAACCAGACATTACTAACAGA 67-8 519 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 67-8 519 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660 501 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 738
Dp	
QY	601 GACACTATTTTTGAAGACTATORD 679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738 679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720 661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCGGATAAAAGT 798
ठंद	TO A COURT A A CTCATGATGAACCTCCTTCGGGGTT
ŷý.	661 CAGTCTTTAAAGCTGCTAGGGGACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798 739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 780 11 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 780 721 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTCACAAA 858
Dp.	721 AAGTATATCAGCATAGATAGATATTAAGGTGTTTTGTGGCCAGTCCTCACAAA 858
Qу	721 AAGTATATCAGCAAGCCGGAGTTTTTTTTTTTTTTTTTT
ďŒ	781 CCCAACATCCAGT2
Qу	781 CCCAACATCCAGTTTGAAGCCTTTO 859 ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 918
Db	841 ACACAGCCTATTGTGGAGAA

The second of th

```
919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
              901 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960
Qy
Db
          979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              QУ
          961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996
Db
RESULT 7
                                                                 ROD 16-APR-2003
                                                        linear
            Mus musculus RIKEN cDNA 1500031K13 gene, mRNA (cDNA clone MGC:28889
 BC016128
 LOCUS
             IMAGE:4911640), complete cds.
 DEFINITION
             BC016128
 ACCESSION
             BC016128.1 GI:16359341
 VERSION
             MGC.
 KEYWORDS
             Mus musculus (house mouse)
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 SOURCE
             Mus musculus
   ORGANISM
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
              Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
  REFERENCE
              Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
              Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
    AUTHORS
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
              Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J..
              Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
              McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
               Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
               Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
               Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
               Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
               Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
               Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
               Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
               Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
               Generation and initial analysis of more than 15,000 full-length
               numan and mouse cDNA sequences
     TITLE
               Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
     OURNAL
                22388257
     MEDLINE
                12477932
       PUBMEL
                   (bases 1 to 1359)
    REFERENCE
                Strausberg, R.
     AUTHORS
                Submitted (22-OCT-2001) National Institutes of Health, Mammalian
                Gene Collection (MGC), Cancer Genomics Office, National Cancer
      TITLE
                Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
      JOURNAL
                NIH-MGC Project URL: http://mgc.nci.nih.gov
      REMARK
                Contact: MGC help desk
                Email: cgapbs-r@mail.nih.gov
    COMMENT
                Tissue Procurement: Jeffrey E. Green, M.D.
                CDNA Library Preparation: Life Technologies, Inc.
```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 38 Row: m Column: 23 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not dentity to protein. Location/Qualifiers 1. .1359 source /organism@"Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:28889 IMAGE:4911640" /tissue_type="Salivary gland, 10 week old remale mouse" /clone_lib="NCI_CGAP_SG2" /lab_host="DH10R" /note="Vector: pCMV-SPORT6" 1. .1359 /gene="1500031K13Rik" /note="synonyms: 4930520C08Rik, 2810425013Rik" /db_xref="LocusID:69008" /db_xref="MGI:1916258" 262. .1266 /codon start=1 /product="1500031K13Rik protein" /protein_id="AAH16128.1" /db_xref="GI:16359342" /db_xref="LocusID:69008" /translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ AMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQ1FNNIL RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN YVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP 346 t 294 g 418 a 301 C BASE COUNT Score 860.4; DB 10; Length 1359; 84.9%; Query Mauch Pest Local Similarity 90.5%; Pred. No. 3.2e-200; Gaps Indels 96; 0; Mismatches Matches 918; Conservative 1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60

QУ

ORIGIN.

FRATURES

ത്രദ്

-1	253 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 312
Db	TOTAL A A GGA A CA CA A A A A GA CAGA CAAGGCTTCAGAA 120
Qγ	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAATTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTO
Db	A CANATTOTGTGTGTGCAAACGAGAAAGAA 180
QУ	121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGTG
Db	THE REAL GEOGREPH CONTROL OF THE REAL PROPERTY OF T
QУ	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTA
Db	THE CASE OF THE TOTAL CONTROL AND ADDRESS OF THE CASE
QУ	241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAATTTTTTTT
Db	THE GRADIEN GOOD CTICGGAGTCCTACTGTGGAGTATATTAGT 360
Őλ	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCGTTTTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCGTTTTTTTT
Dic	A CONTATON A CONTATON A CCCCCACAGATTGCCTTA 420
Qy Db	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCGGTGTTTTTTATGCTCCTCAAAGGATATGAAGCCGGTGTTTTTTTT
	GA GA AUGUA TTCCA CATGA ACCACTTGCCAAAATCATCCTC 480
Ğλ	421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGMC531011111111111111111111111111111111111
Db	GCTGTGGGATTATGCTAAGAGASTSTTCTCAACATTTGATATTGCT 540
ing T	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
2.833	COLOR WITTO CITA ACCAGA CATAAAGTGTTGGTAGCAGAC 600
S. S.	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACTTTACTAACCAGACTTTACTAACCAGACTTTACTAACCAGACTTTACTAACCAGACTTTACTAACCAGACTTTACTTAACCAGACTTTAAAGTACTAGCAGACAGA
DL:	THE STATE OF THE S
Őλ	691 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTGAGAAATTGCTTCAGTGAGAAATTGCTTCAGTGAGAAATTGCTTCAGTGAGAAATTGCTTCAGTGAGAAATTGAGAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 912
Db	853 TTCTTAGAACAAAATTATGACACTTTTTTTTTTTTTTTT
Ç.Y	661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Db	721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
Qŷ	721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACGTATATCAGCAAGCCGGAGAACGTATATGACCAAGCCAGAGCAAACGTATATCAGCAAGCCAGAGAACCTGAAACTGATGAAC 1032 973 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1032
Dþ	781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
ÖÄ	781 CTCCTTCGGGATAAAGTCCCAACATCCAGTTTGAAGCCTTCTTTGATGTCTTTTTTTT
Db	841 GTGGCCAGTCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Qy	841 GTGGCCAGTCCTCACAAACACAGCCTATTGTGGACATCTTTTTTTT
Db	1093 GTGGCCAGCCCCACATTILOGGST-1-

```
Qу
         1153 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGACGACGACGAGCAGCTTTGCTGAC 1212
Db
          961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              1111111111 | 111111111111 | 11111 | 11111 | 11111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 1111 | 11
QУ
         1213 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1266
Db
RESULT 8
                                                                 ROD 16-APR-2003
                                                        Tinear
BC016546
                                                mRNA
                                     1530 bp
            Mus musculus RIKEN cDNA 1500031K13 gene, mRNA (cDNA clone MGC:27972
LOCUS
DEFINITION
             IMAGE:3595339), complete cds.
             BC016546
ACCESSION
             BC016546.1 GI:16741456
VERSION
 KEYWORDS
             MGC.
             Mus musculus (house mouse)
 SOURCE
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mus musculus
   CRGANISM
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
             1 (bases 1 to 1530)
             Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 PEFERENCE
             Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
   AUTHORS
             Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
             Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
             Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Scheetz, T.E., Brownstein, M.J., Usdin, T.B.. Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
              Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
              McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
              Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
              Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, K., Gibbs, R.A.,
              Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
              Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
              Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
              Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
              Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
              Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
              Generation and initial analysis of more than 15,000 full-length
    TITLE
              human and mouse cDNA sequences
               Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
    JOURNAL
               22388257
    MEDLINE
               12477932
      PUBMED
               2 (bases 1 to 1530)
   REFERENCE
               Strausberg, R.
     AUTHORS
               Submitted (31-OCT-2001) National Institutes of Health, Mammalian
     TITLE
               Gene Collection (MGC), Cancer Genomics Office, National Cancer
     JOURNAL
               Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
               NIH-MGC Project URL: http://mgc.nci.nih.gov
     REMARK
               Contact: MGC help desk
   COMMENT
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Jeffrey Green M.D.
```

```
CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          ENA. Sequencing by: Baylor College of Medicine Human Genome
          Sequencing Center
          Center code: BCM-HGSC
          .Web site: http://www.hgsc.bcm.tmc.edu/cdna/
           Contact: amg@bcm.tmc.edu
           Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
           Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
           A.N., Gibbs, R.A.
           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series: IRAK Plate: 35 Row: m Column: 15
           This clone was selected for full length sequencing because it
           passed the following selection criteria: Similarity but not
           identity to protein.
                    Location/Qualifiers
FEATURES
                    1. .1530
    source
                    /organism="Mus musculus"
                    /mol type="mRNA"
                    /strain="FVB/N"
                    /db_xref="taxon:10090"
                    /clone="MGC:27972 IMAGE:3595339"
                     /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
                     ductal carcinoma. 5 month old virgin mouse."
                     /clone_lib="NCI_CGAP_Mam6"
                     /lab host="DH10B"
                     /note="Vector: pCMV-SPORT6".
                     1. .1530
     sene
                     /gene="1500031K13Rik"
                     /note="synonyms: 4930520C08Rik, 2810425013Rik"
                     /db_xref="LocusID:69008"
                     /db_xref="MGI:1916258"
                     279. .1283
                     /codon_start=1
                     /product="1500031K13Rik protein"
                     /protein id="AAH16546.1"
                     /db_xref="GI:16741457"
                      /db_xref="LocusID:69008"
                      -translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ
                      {\tt AMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL}
                      RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
                      QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN
                      YVTKRQSLKLRGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
                      FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
                                   326 g
                                            393 t
                          313 C
                 498 a
 DASE COUNT
 ORIGIN
                                   Score 858.8; DB 10; Length 1530;
                           84.7%;
   Query Match
                                   Pred. No. 8e-200;
                           90.4%;
   Bost Local Similarity
                                                                      Gaps
                                                        Indels
                                  0; Mismatches
                                                   97;
   Matches 917; Conservative
             1 ATCAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
                270 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 329
```

QУ

Dio

QΫ		CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC
Db		CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC
Qÿ		GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGTG
Db		GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGTG
Qy Db		CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
		ACACTGATAGCTGACCTGCACCTGATAGACT'ITGAGGGAAAAAAAGATGTGACCCAGATA 300
Db		ACACTGATAGCTATAGCTATAGAGTGAGGGAAAAAAAAAA
QУ		TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Db	570	TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 629
QУ		GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
ďű		TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 689
Qy Sh		CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 4802
		THE TOTAL A TECA ATTICACA CA CATTETT A AGTA CGTGGAGTTGTCAACATTTGATATTGCT 540
12y - Dâ		TTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 809
QY.	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600;
DE	ខ្វាំប	TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 869
Qy		TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 560
Db		TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 929
Qy	· .	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 72.0-
Db.) AACTATGTGACAAAGAGACAATCTTTAAAGTTGCGAGGTGAGCTGATCCTGGACCGCCAC 989
Çy		AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780.
Db		D AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1049
δλ		CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Db		O CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1109
Qу		1 GTGGCCAGTCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
CC	111	O GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1169

.

```
901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
QУ
             1176 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGACGACGACGAGCAGCTTTGCTGAC 1229
Db
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QУ
             311141141141 131411111111 114141111111 1 41111 1 41111
         1230 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1283
Db
RESULT 9
BD147463
                                                              PAT 17-JAN-2003
                                                      linear
                                              DNA
                                    822 bp
            BD147463
LOCUS
           Primer for synthesizing full-length cDNA and use thereof.
DEFINITION
            BD147463
ACCESSION
            BDi47463.1 GI:27853221
VERSION
            JP 2002191363-A/2306.
KEYWORDS
            Homo sapiens (human)
SOURCE
            Homo sapiens
  ORGANISM
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases I to 822)
REFERENCE
            Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.
  AUTHORS
            Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
            Primer for synthesizing full-length cDNA and use thereof
  TITLE
            Patent: JP 2002191363-A 2306 09-JUL-2002;
  JOURNAL
            HELTK RESEARCH INSTITUTE
                Homo sapiens (human)
COMMEND
                 JP 2002191363-A/2306
            \sim NT
                 09-JUL-2002
            PD
                 28-JUL-2000 JP 2000280990
            DF.
                 TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
            PI.
            ΡI.
               SAITO,
                 JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
            ΣĊ
            PI KEIICHI NAGAI, TETSUJI OTSUKI
            C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
             10.
                 C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
            РC
             Primer for synthesizing full-length cDNA and use thereof FH
                       Location/Qualifiers
            FT
                 source
                                 /organism='Homo sapiens (human)'.
            \mathbb{F}\mathrm{T}
                     Location/Qualifiers
 FEATURES
                     1. .822
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                                                      3 others
                                           216 t
                                  171 g
                268 a
 BASE COUNT
 ORIGIN
                          76.0%; Score 770.6; DB 6; Length 822;
   Query Match
                          98.5%; Pred. No. 3.4e-178;
   Best Local Similarity
                                                                    Gaps 2:
                                                       Indels
                                 0; Mismatches
                                                  10;
   Matches 798; Conservative
            19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
 Qy:
               1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
 Db
```

QΥ	7) ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138	3
Db	7.) ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTGAGAAGAAGACAAAGGCAAGACAAGGCTTGAGAAGAAGAAAGCAAAAAGACAAAAAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120	
Qy	THE	
Db	139 CAAGCAATGAAAGAAATTCTGTGTGTACAAACGAGATTCTAGTGTACAAACGAGATTCTGTGTGTG	
Qу	199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 25	
Db	181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGAAACTCTAAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGAAACTCTACAGCAGTGGCCTGCTGGTGACACTGAAACTCTAAAACTCTACAGCAGTGGCCTGCTGGTGACACTGAAACTCTAAAACTCTACAGCAGTGGCCTGCTGGTGACACTGAAACTCTAAAACTCTACAGCAGTGGCCTGCTGGTGACACTGAAACTCTAAAACTCTAAAACTCTACAGCAGTGGCCTGCTGGTGACAACTACTAAAACTCTAAAAACTACTAAAAAACTCTAAAAAA	
Qу	259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 31	
Dp	241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTMACAAGATGTGAGACCCAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGATATTTMACAAGATGTGAGAGAGATATTTMACAAGATGTGAGAGAGAAGAAGAAGAAGATGTGAGAGAGA	
QΥ	319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 37	
Db	301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATAAGTGGTGAT	
ďÃ,	3.79 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 43	
Dp	A STANCARCA A SCACTEGGGAAAATCATCCTCTTTTCTAATCAATTCAGA 4	
20 <u>7</u>	AGAGAATGTATTCGACATGAACCACTTGCCAAATTCTTTTTTTT	
	GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 5	5.8 4.0a
Ωγ' Db	TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 6	
OD OV	619 GACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 6 	78 60
Qy Db	679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 7	38
	TO THE TRANSPORT OF A CONTROL OF A CONTROL OF THE TRANSPORT OF THE TRANSPO	
Db	73 AAGTATATCAGCAAGCCGGAGAACCTG	
Qÿ	799 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828	
Db	779 CCCAACATCCAGTTTGAGCCTTCTGGTTTT 808	

RESULT 10 BD079551 LOCUS

BD079551

```
DEFINITION Cancer-associated nucleic acids and polypeptides.
          BD079551
ACCESSION
          BD079551.1 GI:22625154
VERSION
          JP 2001516009-A/217.
KEYWORDS
          Homo sapiens (human)
SOURCE
          Homo sapiens
 ORGANISM
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 831)
REFERENCE
          Old, L.J., Scanlan, M.J., Stockert, E., Gure, A., Chen, Y.T., Gout, I.,
 AUTHORS
          OGhare, M., Obata, Y., Pfreundschuh, M., Tureci, O. and Sahin, U.
          Cancer-associated nucleic acids and polypeptides
 TITLE '
          Patent: JP 2001516009-A 217 25-SEP-2001;
  JOURNAL
          LUDWIG INSTITUTE FOR CANCER RESEARCH
              Homo sapiens (human)
COMMENT
              JP 2001516009-A/217
          PN
              25-SEP-2001
              15-JUL-1998 JP 2000503425
          PF
                                                     60/061599 PR
                              08/896164,10-OCT-1997 US
              17-JUL-1997 US
          PR
                         60/061765,10-OCT-1997 US
                                                 08/948705 PR
          10-OCT-1997 US
                                                              LLOYD
                                                 09/102322 PI
                         9721697.2,22-JUN-1998 US
          11-OCT-1997 GB
          J OLD, MATTHEW J SCANLAN, ELISABETH STOCKERT, ALI GURE, YAO PI
          CHEN.
              IVAN GOUT, MICHAEL O'HARE, YUICHI OBATA, MICHAEL PFREUNDSCHUH, PI
            OZLEM TURECI,
              UGUR SAHIN
          PI
          ₽C
          001N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC 3
           A61P35/00,
             C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,
          DC
               C12N15/00
           PC
               Cancer-associated nucleic acids and polypeptides. FH Key
                 Location/Qualifiers
                            1. .831
          TT
               source
                            /organism='Homo sapiens (human)'.
           FT
                  Location/Qualifiers
PERTURES
                  1. .831
     source
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                                     209 t
                                              5 others
                             167 g
                      165 c
              285 a
BASE COUNT
 CRIGIN
                                               Length 831;
                             Score 684.6; DB 6;
                       67.5%;
  Query Match
                      96.1%; Pred. No. 4.1e-157;
  Best Local Similarity
                                                           Gaps
                            0; Mismatches
                                               Indels
         764; Conservative
           1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
 ÕУ
             37 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 96
 ď
          Çy
             Db
         QУ
```

Db	157 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
Çу	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Db	217 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 270
Qу	241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Db	277 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 336
Qy	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Db	301 TITAACAACATCTTGAGAAGACATCTTGTGGAGTCCTACTGTGGAGTATTAGT 396 337 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 396
Óλ	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Db	397 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 456
ÇY	421. CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCAC''TTGCCAAAATCATCCTC 480
Db	421. CGTTGTGGGATTATGCTGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 516
QУ	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
Db	1111CTAATCAATTCAOAGTTTTCTTTTTTTTTTTTTTTTTTTT
Qy	562 ECAGATGCCTTTGCTACTTTCAA-GGATTTACTAACCAGACATAAAGTGTTGGTAGC-AG 598
da da	TCAGATGCCTTTGCTACTTTCAAGGGATTTACTAACCAGACATAAAGTGTTGGTAGCAAG 636
e e e e e e e e e e e e e e e e e e e	ACTITCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 658
Db	637 ACTTCTTAGAACAAAATTACGACACTANTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 696
Óλ	659 AG-AATTATGTTACTAAGAGACAGTCTTTAAAG-CTGCTAGGGGAGCTGATCCTGGACCG 716
Db	697 AGAAATTATGTTACCAAGAGACAGTCCTTAAAGCCTGCTAAGGGAACTGATTCTGGACCG 756
Qу	717 TCACAACTTTGCCATC-ATGACAAAG'IATATCAGCAAGCCGGAGAACCTGAAACTCA 772
DĎ	757 TCANAACTTTGCCATCAANGCAAAAGTTTATCAACAAGCCNGGGGAAACCGGAAACNCAA 816
QУ	773 TGATGAACCTCCTTC 787
Db	817 GGAGGAACCTCCTTC 831
RESULT 1 AX061831	
LOCUS DEFINITI	ON Sequence 1 from Patent WO0078947.
ACCESSIO VERSION	N AX061831 AX061831.1 GI:12539911
KEYWORDS SOURCE	Homo sapiens (human)
ORGANI	

- 0

: :

```
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
          den Daas, I., Fischer, V., Seyfried, C. and von Melchner, L.
 AUTHORS
          Head trauma induced cytoplasmatic calcium binding protein
 TITLE
          Patent: WO 0078947-A 1 28-DEC-2000;
 JOURNAL
          MERCK PATENT GmbH (DE)
                  Location/Qualifiers
FEATURES
                  1. .1026
    source
                  organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db xref="taxon:9606"
                  1. .1026
    COS
                  /note="unnamed protein product"
                  /codon start=1
                  /protein id="CAC25030.1"
                  /db xref="GI:12539912"
                  /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                  KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                  NNILRRQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                  LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                  HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                  VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                   RPAOOEA"
                                      265 t
                      199 c
                              203 g
              359 a
BASE COUNT
ORIGIN
                                                Length 1026;
                              Score 582.6; DB 6;
                       57.5%;
  Query Match
                              Pred. No. 4.3e-132;
                                                             تنجير
  Nest Local Similarity.
                       74.78;
                                                            'Gaps
                             0; Mismatches 244;
                                                Indels
         748; Conservative
 ches
          18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
٠<u>/</u>
             Hili
          12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 71
                                    -AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
          78 CATTTTGGAAAAGCAAGAC-
Qy
                                           1111 11111 11111111111
                                     72 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 131
Db
         111
         132 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 191
 Db
         389 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
 Qу
             152 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 251
OC.
         249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 30,8
                   Qу
         252 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 311
 Db
         300 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
 Ωy
                 31.2 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCAGCCAACA 371
 Db
          369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
 \Omega Y
                    1111 111 1 1 11111 11111 1 111 1 11
                                                        - 11
          372 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 431
```

Db

```
429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
QУ
           432 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 491
Db
       439 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
                  QУ
       492 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 551
Db
       549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
QУ
           552 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 611
Db
       609 ACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
QУ
                             Db
       569 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
Qy.
                      Db
        729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
           QY-
        732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 791
Die
        789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848
           QT
        AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 851
111
        EGG TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908;
           TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 911
DЫ
        909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
\Omega Y
           1111111
        912 GTTCCTCAGCAAGTTTCAGAACGACAGGACGAGGATGAGCAGTTTAACGACGAGAAGAC
        969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
99
           972 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1012
 dG
                                                   . ñ.
 RESULT 12
 M1082322
                                                 PAT 28-FEB-2001
                                          linear
                                    DNA
                          3281 bp
          AX082322
 LOCUS
          Sequence 26 from Patent W00111032.
 DEFINITION
          AX082322
 ACCESSION
          AX082322.1 GI:13184499
 VERSION
 KEYWORDS
          Homo sapiens (human)
 SOURCE
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  ORGANISM
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
          Hodgson, D.M., Lincoln, S.E., Russo, F.D., Spiro, P.A., Banville, S.C.,
          Bratcher, S.R., Dufour, G.E., Cohen, H.J., Rosen, B.H., Chalup, M.S.,
  AUTHORS
          Hillman, J.L., Jones, A.L., Yu, J.Y., Greenawalt, L.B., Panzer, S.R.,
```

```
Roseberry, A.M., Wright, R.J. and Daniels, S.E.
                             Secretory molecules
    TITLE
                              Patent: WO 0111032-A 26 15-FEB-2001;
    JOURNAL
                              Incyte Genomics, Inc. (US)
                                                     Location/Qualifiers
FEATURES
                                                     1. .3281
            source
                                                     /organism="Homo sapiens"
                                                     /mol_type="genomic DNA"
                                                     /db xref="taxon:9606"
                                                     /note="Incyte ID No: 481257.3"
                                                                                                             990 t
                                                               601 C
                                                                                      676 q
                                      1014 a
BASE COUNT
ORIGIN
                                                                                                                                          Length 3281;
                                                                                      Score 582.6; DB 6;
                                                                  57.5%;
     Query Match
                                                                  74.7%;
                                                                                      Pred. No. 4.3e-132;
     Best Local Similarity
                                                                                                                                                                                                   1;
                                                                                                                                           Indels
                                                                                                                                                                     9;
                                                                                                                                                                               Gaps
                                                                                    0; Mismatches 244;
                          748; Conservative
                            18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
 Qy-
                                                                                       1111 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 
                          101 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 160
 Dh
                                                                                                    ---AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
                             78 CATTTTGGAAAAGCAAGAC---
 QУ
                                                                                                                             11111
                                          161 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 220
 Db
                           THE CHARGE CONTROL OF THE STATE OF THE STATE
 QУ
                           CAAAAATCTGGTTGCCATGAAAGAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 280
 137
                           248 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
                                      MARTINE ARREST DESTRUCTION SE ARREST.
                                                                                                                                                                         11 111 1
                           AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 340
  Db
                           308. AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308.
  QУ
                                                       341 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAAGACGTGGCTCAAATTTTCAACAA 400
  Db
                           309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
                                                 Qy
                           401 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 460
  Db
                            369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
                                                          ÇУ
                          461 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 520
  Db
                            429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
   QУ
                                                                                                 11 111 1 1111111
                            521 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 580
   Db
                            439 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
                                                                                                                        QУ
                                                                                            58. ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 640
   Db
                             549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
    Qу
                                          641 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 700
```

Db

意大大

```
609 ACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
                               11 11111111 11 11111 11 11 111111111
Qv
           Dia
        669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
            QУ
        Db
        729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
            Qy
        821 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 880
Db
        789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848
            QY
        881 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTTGTAGCCAA 940
D.b
        849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAACTCATTGA 908
           Qу
        941 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1000
Ľρ
        909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
            Qv
        \Gamma ib
        959 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
            18 11 .111111111 1 11.1111111 1.1 191
 Qу
        2031 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1101
 Db
 11,50% (13.1
                                                       44 . 20.
                                                     PRI 22-JAN-2002
 18020570
                                             linear
                                     . mRNA
                              3761 bp
          Homo sapiens, MO25 protein, clone MGC:21631 IMAGE:4397573, mRNA,
 100125
 DELINITION
           complete cds.
           BC020570
 ACCESSION -
           BC020570.1 GI:18088260
 VERSION '
 KEYWORDS
           MGC.
           Homo sapiens (human)
 SCURCE
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Homo sapiens
   CRGANISM
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 3761)
 REFERENCE
           Strausberg, R.
   AUTHORS
           Submitted (03-JAN-2002) National Institutes of Health, Mammalian
           Direct Submission
   TITLE
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
   JOURNAL
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           NIH-MGC Project URL: http://mgc.nci.nih.gov
   REMARK
           Contact: MGC help desk
  COMMENT
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Life Technologies, Inc.
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Sequencing Group at the Stanford Human Genome
           Center, Stanford University School of Medicine, Stanford, CA 94305
                        http://www-shgc.stanford.edu
```

```
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
          P. M.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 27 Row: d Column: 16
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Hexamer frequency ORF
          analysis, GenomeScan gene prediction.
                  Location/Qualifiers
FEATURES
                  1. .3761
    cource
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="LocusID:51719"
                  /db xref="taxon:9606"
                   /clone="MGC:21631 IMAGE:4397573"
                   tissue type="Duodenum, adenocarcinoma"
                   /clone lib="NIH_MGC_88"
                   /lab_host="DH10B"
                   /note="Vector: pCMV-SPORT6"
                   325. .1350
    ODS.
                   /codon_start=1
                   /product="MO25 protein"
                   /protein_id="AAH20570.1"
                   /db xref="GI:18088261"
                   /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                   KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                   NNILRRQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                   LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                   HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                   VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                   RPACOEA"
                                      1073 t
             1171 a
                      709 C
                              808 g
JASE COUNS
ORIGIN
                              Score 582.6; DB 9;
                                                 Length 3761;
                       57.5%;
  Query Match
                       74.7%; Pred. No. 4.3e-132;
  Best Local Similarity
                              0; Mismatches 244;
                                                 Indels
  Matches 748; Conservative
          18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGC 77
QУ
             336 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGCATGGC 395
Db
                                     -AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
          78 CATTTTGGAAAAGCAAGAC---
QV
                                      396 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAAGGCTACAGAAGAAGTTTC 455
55.
         QУ
                         456 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 515
Db
         189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
 ΩΣ.
             516 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 575
 Db
         249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
 QУ
```

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Db	
	200 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
ДÀ	
Db	369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
ĊХ	696 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 755
Db	696 GAATATTTTGTTCATGTTATTGAAACCOTTTGCCAAAATCATCCTCTTTTCTAA 488
ÕУ	423 GATTATGCTGAGAGATGTATTCGACATGAACCACTTGCATATATGTGTCGAGAGATGTATTCGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 815
Db	\cdot
Qу	489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
Db -	816 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 875
ÇУ	549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
Db	876 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTTGGA 933
ŌХ	609 ACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
Th.	ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCAGTCAG
ΩY	TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 7284
.Dp-	GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTAGATAGA
ijij.	CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788.
DŽ3	1056 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1115
Ger	789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848
Dħ	
ÔΛ	849 TCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
- 7 - da	Tectracaaaacacacacciaficiosia.http://doi.org/10.1016
Qy	909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
Σp	1236 GTTCCTCAGCAGCTTCCAAAAAAAAAAAAAAAAAAAAAA
	969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
Ö.	
Db	1296 CIAITTAGITAAACAGATGAGGATTA

RESULT 14

AF151824

AF151824 LOCUS

1680 bp

mRNA linear PRI 18-MAY-2000

DEFINITION Homo sapiens CGI-66 protein mRNA, complete cds.

ACCESSION

AF151824

```
AF151824.1 GI:4929600
VERSION
KEYWORDS
          Homo sapiens (human)
SOURCE
          Homo sapiens
 ORGANISM
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              (bases 1 to 1680)
REFERENCE
           Lai, C.H., Chou, C.Y., Ch'ang, L.Y., Liu, C.S. and Lin, W.
  AUTHORS
           Identification of novel human genes evolutionarily conserved in
  TITLE
           Caenorhabditis elegans by comparative proteomics
           Genome Res. 10 (5), 703-713 (2000)
  JOURNAL
           20272150
  MEDLINE
           10810093
  PUBMED
           2 (bases 1 to 1680)
REFERENCE
           Lin, W.-C.
  AUTHORS
           Direct Submission
  TITLE
           Submitted (17-MAY-1999) Institute of Biomedical Sciences, Academia
  JOURNAL -
           Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan
                   Location/Qualifiers
FEATURES
                   1. .1680
    source
                   /organism="Homo sapiens"
                   /mol type="mRNA"
                   /db xref="taxon:9606"
                   1. .1026
     CDS
                   /codon start=1
                    /product="CGI-66 protein"
                    /protein_id="AAD34061.1"
                    /db xref="GI:4929601"
                   /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                   KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                   NNILRRQIGTRTPTVEYICTQQNILFMLKGYESPEIALNCGIMLRECIRHEPLAKII
                   LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                   HSENYVTKROSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                   VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDFQFNDEKTYLVKQIRDLK
                    RPACOEA"
                                358 g 458 t
                       324 c
               540 a
BASE COUNT
ORIGIN
                                Score 581.6; DB 9;
                                                   Length 1680;
                        57.4%;
  Query Match
                        74.7%; Pred. No. 7.5e-132;
  Best Local Similarity
                                                                :Gaps
                                                   Indels
                               0; Mismatches 244;
  Matches 747; Conservative
           19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
 Qу
             13 TTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGCT 72
 Db
                                      -AAAAAGACAGACAAGGCTTCAGAAGAAGTGTCT 129
           79 ATTTTGGAAAAGCAAGAC---
 QУ
                                       73 GTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAAGGCTACAGAAGAAGTTTCC 132
 Db
          ÇΥ
                         133 AAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGACA 192
 Db
          190 GAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATA 249
 Çv
              1111111 11111 11 11 11 11 11 11 11 11
          193 GAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGTA 252
 Db
```

Qy		250	GCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAAGATGTGACCCAGATATTTAACAAC	505
Db		253	GCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAAT	312
Çy .		310	ATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATAT'IAGTGCTCATCCT	369
Db		313	ATCTOMENTAL ATTOMES AND ATTOTOGRAPH ATTOTO	372
Qу		3.70	CATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGG	429
ad		373		432
QУ	,	430	ATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT	489
. ojo .		433 3		492
QΫ́		490	CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC	549
Db		493		552
Qу		550	TTTGCTACTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA	609
Db.	-	§ 53	TTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGAA	612
<mark>ዩ</mark> ኔ		₫3. 0	CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT	569]
1251			CAGGATTATOATAGATTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTG	6721
Qys	er e		ACTAAGAGACAGTCTTTAAAGCTGCTACGGGAGCTGATCCTGGACCGTCACAACTTTGCC	72-97
ob Db		£72	CAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTAGATAGA	7.3:27
ÇŢ		730	ATCA IGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGG	789
200		733	HILLIAN STATEMENT OF THE STATEMENT OF TH	792
Q''		798	GATAAAAGICCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGT	849
מַּם		793	GACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAAT	852
-97		35C	CCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAG	909
D'o		. 653		91/2
ŠΣ		.pr	O TTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAAC	969
Db	~	91.		972
 ОУ		57	O TACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009	, to
Db		97		

RESULT 15 AF113536 LOCUS

```
Homo sapiens MO25 protein mRNA, complete cds.
DEFINITION
ACCESSION
           AF113536
           AF113536.1 GI:6523826
VERSION
KEYWORDS
           Homo sapiens (human)
SOURCE
  ORGANISM
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
             (bases 1 to 3466)
REFERENCE
           Jin, W., Shi, J., Ren, S., Gu, J., Fu, S., Huang, Q., Dong, H., Yu, Y.,
  AUTHORS
           Fu,G., Wang,Y., Chen,Z. and Han,Z.
           A novel gene expressed in the human hypothalamus
  TITLE.
           Unpublished
  JOURNAL
           2 (bases 1 to 3466)
REFERENCE
           Jin, W., Shi, J., Ren, S., Gu, J., Fu, S., Huang, Q., Dong, H., Yu, Y.,
  AUTHORS
           Fu,G., Wang,Y., Chen,Z. and Han,Z.
           Direct Submission
  TITLE
           Submitted (16-DEC-1998) Chinese National Human Genome Center at
  JOURNAL
           Shanghai, 351 Guo Shoujing Rd., Zhangjiang Hi-Tech Park, Pudong,
           Shanghai 201203, China
                   Location/Qualifiers
FEATURES
                    1. .3466
     source
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /tissue_type="hypothalamus"
                    54. .1079
     008
                    /codon start=1
                    /product="MO25 protein"
                    /protein id="AAF14873.1"
                    /db_xref="GI:6523827"
                    /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                    KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                    NNILRPQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                    LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                    HSENYVŢKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                    VFKVFVANPNKTQP1LD1LLKNQAKL1EFLSKFQNDRTEDEQFNDEKTYLVKQ1RDLK
                    RPAQQEA"
                                       1070 t
                                689 q
                        606 C
              1101 a
 BASE COUNTY.
 OPIGIN:
                                Score 581.6; DB 9; Length 3466;
                         57.4%;
   Query Match
                         74.7%; Pred. No. 7.6e-132;
   Best Local Similarity
                                                   Indels
                                                                Gaps 1;
                              0; Mismatches 244;
   Matches 747; Conservative
           12 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
 QΥ
              66 TTTGGAAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGCT 125
 Db
                                      -AAAAAGACAGACAAGGCTTCAGAAGAAGTGTCT 129
           73 ATTTTGGAAAAGCAAGAC-----
 QΥ
                                       125 GTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTCC 185
 dd
          QУ
                         111
          186 AAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGACA 245
 Db
```

Qy	190 GAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATA 249
Dp	
Çу	250 GCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAAC 309
Db	310 ATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCT 369
Qy Db	310 ATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGTGTATTAGTGTAGTGTAGATAGA
ממ	370 CATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGG 429
Су Do	
Ov	430 ATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
ДУ	
QУ	490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
Db	
8X	550 TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA 609
Th.	
Ç.Y	CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT 669
5 2	TARGETTE A ACCURACIONA DE CONTROL DE LA CONT
QY .	ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAAGTTATTTTTTTT
Db-	730 ATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGG 789
DP CA	730 ATCATGACAAAGTATATCAGGTATATCAGGTATAGATGATGAACCTGCGAAATTAATGATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCGAAATAAACCTGAGAACCTCAAATTAATGATGAACCTGCGAAATAAACCTGAGAACCTCAAATTAATGATGAACCTGCGAAATAAACCTGAGAACCTCAAATTAATGATGAACCTGCGAAATTAATGAATG
Out	790 GATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGT 849
QY Db	
СУ	850 CCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAG 909
Db	
QY	910 TTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAAC 969
ōb	
δλ	970 TACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
. da	1026 TATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1065

Search completed: January 6, 2004, 02:34:57 Job time: 3971 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model January 6, 2004, 00:31:27; Search time 390 Seconds Run on: (without alignments) 7018.539 Million cell updates/sec

US-10-088-872-1 Title:

1014 Perfect score:

1 atgaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014 Sequence:

IDENTITY NUC Scoring table:

Gapop 10.0 , Gapext 1.0

2552756 seqs, 1349719017 residues Seamphed:

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

N Geneseq 19Jun03:* Datiribade -

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:* /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:* 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:* 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:* 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:* 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:* 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:* 7:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:* /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:* 9: 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:* 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*

12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:* 13:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:* 14:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:* 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*

16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:* 17:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:* 19:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:* 20:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:* 21:

/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:* 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*

23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARĮES

			%				
Resi	ılt		Query				
	TO.	Score		Length	DB	ID	Description
							Human Acute Neuron
	1	1014	100.0	1014	22	AAF86462	Human polynucleoti
	2	1014	100.9	1421	22	AA158234	Human polynucieoti Human calcium bind
	3	1010.8	99.7	1344	21	AAA27332	
	٤Ļ	992.8	97.9	2002	22	AAH15879	Human cDNA sequenc
	5	770.6	76.0	822	22	AAH05471	Human cDNA clone (
	6	684.6	67.5	831	20	AAX39817	Gastric cancer ass
С	7	634.4	67.5	1191	22	AA160020	Human polynucleoti
•	8	582.6	57.5	1026	22	AAC91772	Human ANIC-BP (acu
	9	582.6	57.5	3281	24	ABK13127	Human secretory po
	10	582.6	57.5	3849	23	ABV22987	Human prostate exp
	11	582.6	57.5	3849	23	ABV28822	Human prostate exp
	12	541.6	53.4	1053	22	AAF30688	Human acute neuron
	13	539.6	53.2	1162	23	AAS89557	DNA encoding novel
С	14	520.2	51.3	833	20	AAX39818	Gastric cancer ass
C	15	496	48.9	2492	23	AAS88031	DNA encoding novel
	16	387.8	38.2	722	20	AAZ15133	Human gene express
	1.7	362.8	35.8	2231	23	ABL07151	Drosophila melanog
	18	362.8				ABL07150	Drosophila melanog
	1.9	288.8	28.5			ABS77084	Frog embryonic gen
	20	246.4	24.3			ABL82285	Human ovarian canc
		44.8	24.1			ABL82921	Human ovarian canc
	21. 22	114.3	24.1			ABL81975	Human ovarian canc
		210.3	20.8				Frog embryonic gen
	23	210.4	20.7				Arabidopsis thalia
	24	208.8	20.6				Arabidopsis thalia
	25	200.2	19.7				Arabidopsis thalia
	26	300.2	19.7				Arabidopsis thalia
	27	195	19.2				Gene #481 used to
Ġ.		195	19.2				Lung cancer relate
С	29	169.8	16.7				Porcine muscular s
•	.30	166.6	16.7				Porcine muscular s
	31		16.2				Group III cDNA can
	32	163.8	16.2				Human gene express
	3:3	163.4	15.9				
	34	161.2					Arabidopsis thalia
	35	156	15.4				DNA encoding novel
C	36	153.4					Zea mays DNA fragm
	37	147.2	14.5				T along /
	38	133.2	13.1				lourge parishers and
	39	107.6					lorron wait.
	40	107.6					TO ADDITION OF A CONT
	41	65.6					
	42	65					The state of the s
	43	65					CDD warmage
	44	53.6					
	45	43	4.2	2 44	7 2	1 AACUU443	

```
RESULT 1
AAF86462
     AAF86462 standard; cDNA; 1014 BP.
ID
XX
     AAF86462;
AC
XX
                   (first entry)
     26-JUN-2001
DT'
     Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP, cDNA.
XX.
DE
XX
     Human; cerebroprotective; neuroprotective; vulnerary; vaccine;
     gene therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;
KW
     stroke; acute head trauma; multiple sclerosis; spinal cord injury; ss.
ΚW
KW
XX
      Homo sapiens.
OΞ
XX
                       Location/Qualifiers
FH
      Key
                       1..1014
FT
      CDS
                       /*tag=a
                       /product= "Human Acute Neuronal Induced Calcium Binding
FT
FT
                       Protein, ANIC-BP"
FT
\chi\chi
      W0200123552-A1.
 DM.
 XX
      05-APR-2001.
CT
 XX
      18-8EP-2000; 2000WO-EP09132.
 PV
 \Sigma_{-}X
                      99EP-0118848.
 \Sigma \mathfrak{R}
      _4-SEP-1999;
 \mathbf{X}_{i} \times
       (MIRR ) MERCK PATENT GMBH.
 \mathbb{Z}_{\overline{A}}
 XX
                    Duecker K;
      Dem Daas I,
 ЭI
 XX
      WPT: 2001-308142/32.
 DR
      P-PSDB; AAB82090.
 \mathfrak{DR}
      Novel human acute neuronal induced calcium binding polypeptide; and
 XX
      polynucleotides encoding them useful for diagnosing or treating stroke,
 TT.
 PT
      acute head trauma, multiple sclerosis and spinal cord injury -
 \mathbf{p}\mathbf{T}
 XX
       Claim 5; Page 40-41; 45pp; English.
 PS
       The present sequence is the coding sequence for human Acute Neuronals.
 ХX
       Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and
 CC
       protein are useful for treating stroke, acute head trauma, multiple
 CC:
       sclerosis and spinal cord injury. ANIC-BP coding sequence and protein
 CC:
       are also useful as vaccines for inducing an immunological response in a
 CC
 CC
       mammarl.
 CC
 XX
       Sequence 1014 BP: 340 A; 205 C; 209 G; 260 T; 0 other;
  SO
                                                              Length 1014;
                              100.0%; Score 1014; DB 22;
    Query Match
                              100.0%; Pred. No. 3.5e-272;
    Best Local Similarity
                                                                                     0;
    Matches 1014; Conservative 0; Mismatches
                                                                            Gaps
                                                        0; Indels
```

Qy	1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Db .	
Qy	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC
Db	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAAAAAGACAAAAAAGACAAAAAA
Qy	121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
Db	
Qy	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Db	TO A GORGON COTTON CACTUTGAGGGAAAAAAAAGATGTGACCCAGATA 300
Дb	241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGGAAAAAAAA
'Qy	TOTAL GARGAGA GACATACCCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Db	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATTAGT 360 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Qy	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
වර්	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCCAAAGGATATGAAGCCCCAAAGGATATGAAGCCCCCAAAGGATATGAAGCCCCCAAAGGATATGAAGCCCCCAAAGGATATGAAGCCCCCAAAGGATATGAAGCCCCCAAAGGATATGAAGCCCCCAAAGGATATGAAAGCCCCCAAAGGATATGAAAGCCCCCAAAGGATATGAAAGCCCCCAAAGGATATGAAAGCCCCCAAAGGATATGAAAGCCCCCAAAGGATATGAAAGCCCCCAAAAGGATATGAAAGCCCCCAAAAGGATATGAAAGCCCCCAAAAGGATATGAAAGCCCCCAAAAGGATATGAAAGCCCCCAAAAGGATATGAAAGCCCCCAAAAGGATATGAAAGCAAAAGGATATGAAAGCCCCCAAAAGGATATGAAAGGATATGAAAGCCCCCAAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGGAATATGAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGGAATATGAAAGCAAAGAAAG
Qγ	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
Est.	
Qy Db	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
	THE THE COURT OF THE CANCES THE COURT OF THE
Qy Db	541 TCAGATGCCTTTGCTACTTCAAGGATTTACTMCTATTCTTCAAGGATTTACTMCTATTCTTTCAAGGATTTACTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600.
, LAD	691 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
Qy E/b	601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACATATGAGAAATTGCTTCAGTCTGAG 660
	720
ĆΥ	AATTATGTTACTAAGAGACTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTO
Do /*	721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
QУ	721 AACTTTGCCATCATGACAAAGTATATCAGGTTGOODIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Db	781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAACGTGTTT 840
Qy	781 CTCCTTCGGGATAAAAGTCCCAACATGATTTTTTTTTTT
Db Ov	341 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
QУ	

*1

```
841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Db
          901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
              QУ
          901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Db
          961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              QУ
          961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Db
RESULT 2
AAI58234
     AAI58234 standard; cDNA; 1421 BP.
TD
XX
     AAI58234;
AC
XX
     22-0CT-2001 (first entry)
DT
ХX
     Human polynucleotide SEQ ID NO 437.
DΕ
     Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX
     peripheral nervous system; neuropathy; central nervous system; CNS;
ΧW
     Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 \mathbb{K} \mathbb{W}
     amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 \mathbb{Z}\mathbb{N}
      chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 W_{2}^{\infty}
 ŔЧ
      loukaemia; ss.
 20!
 XZ
      Mosko sapiens.
 OS
             4 4
 XX
      7:0200153312-A1.
 PM
 XX
      26-JUL-2001.
 223
 XX
      25-DEC-2000; 2000WO-US34263.
 PF
 XX
      21-JAN-2000; 2000US-0488725.
 ЭR
      25-APR-2000; 2000US-0552317.
 \mathbb{R}
      09-JUL-2000; 2000US-0598042.
 \mathbf{p}\mathbf{p}
      19-JUL-2000; 2000US-0620312.
 PR
      03-AUG-2000; 2000US-0653450.
 \mathbb{P}\mathbb{R}
      4-SEP-2000; 2000US-0662191.
 . DR
      19-OCT-2000; 2000US-0693036.
 PR
      29 NOV-2000; 2000US-0727344.
 PR
 XX
       (HYSE-) HYSEQ INC.
 PΑ
      Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 XX
                                                                      Wang D;
 PI
 PΙ
                Zhou P, Goodrich R, Drmanac RT;
 ΣŢ
       Zhao QA,
 XX
       WPI; 2001-442253/47.
 \overline{DR}
       P-PSDB; AAM39078.
  DR
  XX
       Novel nucleic acids and polypeptides, useful for treating disorders
  PT
       such as central nervous system injuries
  PT
```

```
XX
   Claim 1; SEQ ID NO 437; 10078pp; English.
PS
XX
   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC
   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC
   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC
    in gene therapy. A composition containing a polypeptide or polynucleotide
CC
   of the invention may be used to treat diseases of the peripheral nervous
CC
   system, such as peripheral nervous injuries, peripheral neuropathy and
CC
    localised neuropathies and central nervous system diseases, such as
CC
   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC
    lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC
    utilisation of the activities such as: Immune system suppression,
CC
    Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC
    and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC
    assays for receptor activity, arthritis and inflammation, leukaemias and
CC
    C.N.S disorders.
CC
    Note: The sequence data for this patent did not form part of the printed
CC
    specification.
CC
XX
    Sequence 1421 BP; 469 A; 284 C; 306 G; 362 T; 0 other;
SÓ
                          Score 1014; DB 22; Length 1421;
                    100.0%;
 Query Match
                    100.0%;
                          Pred. No. 4e-272;
  Best Local Similarity
                         0; Mismatches
                                          Indels
                                                     Gaps...
                                       0:
  Matches 1014; Conservative
         1 ATGAAAAAAATGCCTTTGTTAGTAAATCACAGAAAAATCCAGCAGAAATTGTGAAAATC 60
Qy I
           ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276
:/b
        够/
           ו',כ
        \mathfrak{D}\lambda
           Db
        381 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
ΩУ
           CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
Db
        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
QУ
           457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516
Db
        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Qy
           517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
Db
        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
QΥ
           577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
Db
        421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
Qÿ
           637 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 696
```

Db

```
481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           QУ
        697 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 756
Db
        \pm 4.1 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC \pm 6.00
           QУ
        757 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 816
Db
        601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
           Qy
        817 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 876
Db
        561 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
           QУ
        877 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 936
Db
        721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
           QУ
        937 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 996
OO
        781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
            Qу
        997 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 1056
        900 GET GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA
            .057 OTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAA 1116
        Dell CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 9600.
            ANDRIGHTÉ Í HÁTHARÐI FRANSKAURSKA Í FRÁÐRAÐI
        LL17 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1176
Db
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 ДĀ
            1177 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1230
 Db
 RESULT 3
 AAA27332
     PAA27332 standard; cDNA; 1344 BP.
 ID
 XX
 AC
     MAA27332;
 XX
               (first entry)
     20-AUG-2000
 DT
 XX
     Mussan calcium binding protein hCBP gene.
 DE
 ХX
     Hur,an; calcium binding protein; cancer; inflammation; CBP;
     reproductive disorder; autoimmune disorder; developmental disorder;
 KW
 KW
     seizure disorder; immune disorder; infection; ss.
 KW
 XX
     Homo sapiens.
 OS
 XX
                  Location/Qualifiers
 FH
     Kev
                  124..1134
     CDS
 FT
                  /*tag= a
 FT
```

```
/product= "calcium binding protein"
FT
XX
    WO20C029580-A1.
PN
XX
    25-MAY-2000.
PD
XX
                 99WO-US27027.
    12-NOV-1999;
PF
XX
                 98US-0190965.
    13-NOV-1998;
PR
XX
    (INCY-) INCYTE PHARM INC.
PA
XX
    Tang YT, Guegler KJ, Corley NC, Gorgone GA;
PΙ
XX
    WPI; 2000-387793/33.
DR
    P-PSDB; AAY94247.
DR
    Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
XX
    diagnosis, prevention and treatment of cancers, immune, developmental
PT
PT
    or reproductive disorders -
PΤ
XX
    Claim 9; Fig 1; 72pp; English.
PS
    The present sequence is the human calcium binding protein hCBP gene. It
XX
CC
    was obtained by screening a coronary artery smooth muscle cDNA library,
CC
    from which five overlapping nucleic acids were isolated and
     sequenced, and then expressed to give the protein. The protein and the
CC
    gene encoding it are useful for the diagnosis and treatment of the
CC
CC
     following types of disorder: cancers (such as adenocarcinomas);
-
     reproductive disorders (such as infertility, ovulatory defects;
C_{i,i,j}
     endometriosis, disruptions of the oestrus and menstrual cycles;
     polycystic ovary syndrome and ovarian hyperstimulation), autoimmune
CC
     disorders (such as benign prostatic hyperplasia and prostatitis),
CC.
     developmental disorders (such as Cushing's syndrome, muscular dystrophy
CC
     and gonadal dysgenesis), hereditary neuropathies, seizure disorders,
 CC.
 CO
     immune disorders (such as AIDS, allergies, anaemia, asthma,
     atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'.
 CC
 CC
     disease, multiple sclerosis, psoriasis, rheumatoid arthritis,
     scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,
 CC
 CC
     pacterial, fungal, parasitic, protozoal and helminthic infections.
 CC
 XX
     Sequence 1344 BP; 450 A; 261 C; 280 G; 353 T; 0 other;
 SO
                              Score 1010.8; DB 21; Length 1344;
                       99.7%;
   Query Match
                              Pred. No. 3.1e-271;
                       99.8%;
   Best Local Similarity
                                                                   0;
                                                         0; Gaps
                             0; Mismatches
                                               Indels
   Matches 1012; Conservative
           1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
             Qу
         124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 183
 Db
           QУ
          ^{\mathrm{Db}}
          QY
```

Db	244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAAGAA 303
Ολ	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240°
Db Qy	241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Dp Qy	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Db	4.2.4 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATTTATGCTTCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Db Db	
Qy Db	421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
Çy Do	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540 [
Ωy	TCAGATGCCTTTGCTACTTTCAAGGAFTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
G(A	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660;
ÐÞ. Ö}√	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
QΥ	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC
QY Db	731 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Dh	904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT
Qy Db	GTGGCCAGTCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Qy Db	901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
QУ	961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Db	1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137

, ,

e gen

```
RESULT 4
AAH15879
     AAH15879 standard; cDNA; 2002 BP.
ID
XX
     AAH15879;
AC
XX
                   (first entry)
     26-JUN-2001
DT
XX.
     Human cDNA sequence SEQ ID NO:14407.
DE
     Human: primer; detection; diagnosis; antisense therapy; gene therapy;
XX
ΚW
XX
     Homo sapiens.
OS
XX
      EP1074617-A2.
PN
XX
      GT-FEB-2001.
PD
XX
      20-JUL-2000; 2000EP-0116126.
 PF
 XX
                     99JP-0248036.
      99-JUL-1999;
 Ţ:R
                     99JP-0300253.
      7-AUG-1999;
 \delta S
      11-DAN-2000; 2000JP-0118776.
 - R
      MAY-2000; 2000JP-0183767.
 £R
      09-JUN 2000, 2000JP-0241899.
 PR.
 XX
      (HEBI-) RELIX RES INST.
 المانية
المحادية
      ියෙ ්රි, Isogai T, Nishikawa T, Hayashi K, Saito K,
                                                              Yamamoto 🕠;
                Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
      ishii S
 N.S.
      WPT; 2001-318749/34.
 DE.
      Primer sets for synthesizing polynucleotides, particularly the 5602
 XX
      full-length cDNAs defined in the specification, and for the detection
 \mathbf{r}
      and/or diagnosis of the abnormality of the proteins encoded by the
 PT
 PT
       full-length cDNAs -
 PT
      Claim 8; SEQ ID 14407; 2537pp + CD ROM; English.
 XX
  PS
       The present invention describes primer sets for synthesising 5602
 \mathbb{X}\mathbb{X}
       full-length cDNAs defined in the specification. Where a primer set
  CC
       comprises: (a) an oligo-dT primer and an oligonucleotide complementary
  CC
       to the complementary strand of a polynucleotide which comprises one of
  CC
       the 5602 nucleotide sequences defined in the specification, where the
  CC
       oligonucleotide comprises at least 15 nucleotides; or (b) a combination
  CC
       of an oligonucleotide comprising a sequence complementary to the
  CC
       complementary strand of a polynucleotide which comprises a 5'-end
  CC
       sequence and an oligonucleotide comprising a sequence complementary to a
  CC
  CC
       polynucleotide which comprises a 3'-end sequence, where the
       oligonucleotide comprises at least 15 nucleotides and the combination of
  CC
       the 5'-end sequence/3'-end sequence is selected from those defined in
  CC
       the specification. The primer sets can be used in antisense therapy and
  CC
       in gene therapy. The primers are useful for synthesising polynucleotides,
  CC
       particularly full-length cDNAs. The primers are also useful for the
  CC
  CC
```

the full-length cDNAs. The primers allow obtaining of the full-length CDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 CC represent oligonucleotides, all of which are used in the exemplification CC CC of the present invention. CC XXSequence 2002 BP; 594 A; 418 C; 463 G; 527 T; 0 other; SQ Length 2002; Score 992.8; DB 22; 97.9%; Query Match Pred. No. 3.8e-266; 99.8%; Best Local Similarity 0; Gaps Indels Mismatches Matches 994; Conservative 0: 19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78 ÔУ 1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60 Db 79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138 QУ 51 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120 Db 139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 198 Ωy CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 180 Db 199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258 131 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240 1.05 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318 (QY CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300, \Box REACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 373 301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360 Db 379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438 QV361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420 Db 439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498 () y 480 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480 Dis 49.9 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558 Q_T^{ω} 481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540 Db 559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618 ुरः 541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600 Db 678 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA ÇΥ 601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

detection and/or diagnosis of the abnormality of the proteins encoded by

CC

CC

Db

```
679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
QУ
            661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720
Db
         739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
. Qу.
            721 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 780
Db
         799 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAA 858
Qу
            781 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAA 840
Db
         359 ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 918
QУ
            341 ACACAGCCTATTGTGGAGATGCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 900
Db
         919 AGCTTCCAAAAAGAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
 Qу
            901. AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960
 Db
         979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 Qy
             961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996
 Db
 RÉSULT 5.
 AME08471
     WMH05471 standard; cDNA; 822 BP.
 XX
 AC.
     AAH05471;
 ΧX
     36-JUN-2001 (first entry)
 DT
 ХX
     Human cDNA clone (5'-primer) SEQ ID NO:2306.
 DE
 XX
     Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW
 XX
     Momo sapiens.
 OS
 XX
 РŊ
     EP1074617-A2.
 XX
     07-FEB-2001.
 PD
 XX
     Э3-JПL-2000; 2000EP-0116126.
 \mathbf{p}_{\mathbf{F}}
 XX
     29-JUL-1999;
                  99JP-0248036.
 PR
                  99JP-0300253.
     27-AUG-1999;
 PR
     11-JAN-2000; 2000JP-0118776.
 PR
     02-MAY-2000; 2000JP-0183767.
 PR
     09-JUN-2000; 2000JP-0241899.
 PP
 XX
      (HELI-) HELIX RES INST.
 D\Delta
 XX
     Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K,
                                                    Yamamoto J;
 PI
      Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI
```

XX

DR WPI; 2001-318749/34.

XX PT

PT

PT

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

PT XX PS

Claim 1; SEQ ID 2306; 2537pp + CD ROM; English.

CC CC

CC

CC

CC

CC

CC

CC

CC CC

CC

CC

CC

CC

C.C.

70

CC

: !('

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

CC XX SO

Caquence 822 BP; 268 A; 164 C; 171 G; 216 T; 3 other;

- 79 ATTTTGGAAAAGCAAGACAAAAGACAGACAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
- Db 61 ATTTGGAAAAGCAAAAAAGACAGACAAGGCTTCAGAAGAGTGTCTAAATCACTG 120
- Qy CAAGCAATGAAAGAAA1TCTGTGTGTGCTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 198
- Db 121 CAAGCAATGAAAGAAATTCTGTGTGTGTGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 180
- Qy 199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
- Db 181 GCTCAGCTAGCACAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240
- 259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
- Db 241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

```
319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
          QУ
       301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360
Db
       379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
           QУ
       361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420
Db
       439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
Qу
           AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480
Db
       429 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
           QУ
       481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540
Db
        559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
           QУ
        541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600
Lb
        619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
           QУ
        601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660
Db
           CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
           CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTT'IGCCATCATGACA 729
        709 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
           AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 778
200
        759 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828
QY
           1 1.1
        779 CCCAACATCCAGTTTGAGCCTTCTGGTTTT 808
Db
RESULT 6
AAX39817
    AAX39817 standard; DNA; 831 BP.
ΞĐ
XX
AC
    AAX39817;
XX
    02-JUL-1999 (first entry)
DT
XX
    Castric cancer associated gene.
 DΕ
 XX
    Cancer associated antigen; diagnosis; research; treatment; human;
 KW
    breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 YV
     prostate cancer; ss.
 KM
 XX
     Homo sapiens.
 0S
 XX
     W09904265-A2.
 PN
 XX
```

28-JAN-1999.

PD

```
XX
                98WO-US14679.
    15-JUL-1998;
PF
XX
                98US-0102322.
    22-JUN-1998;
PR
                97US-0896164.
    17-JUL-1997;
PR
                97US-0061599.
    10-OCT-1997;
PR
                97US-0061765.
    10-OCT-1997;
PR
                97US-0948705.
    10 OCT-1997;
PR
                97GB-0021697.
    11-OCT-1997;
PR
XX
    (LUDW-) LUDWIG INST CANCER RES.
PA
XX
    Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
Τσ
    Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
PI
    Tureci 0;
DI
XX
    WPI; 1999-132448/11.
DR
    New isolated cancer associated nucleic acids and polypeptides -
XX
    isolated using sera from cancer patients, used to develop products
PT
PΤ
    for the diagnosis, monitoring or treatment of cancers
PT
XX
    Chaim 67; Page 558-559; 787pp; English.
⊋S.
    The invention relates to a method for diagnosing a disorder characterised
XX
    by expression of a human cancer associated antigen precursor coded for by
CC
CO
    a nucleic acid molecule (NAM). The method comprises: (a) contacting, a
    biological sample isolated from a subject with an agent that specifically
03
    binds to the NAM, an expression product or a fragment of an expression
CO
OC
    product complexed with an HLA molecule; and (b) determining the
    interaction between the agent and the NAM or the expression product as a
CC
CC.
    determination of the disorder. The products and methods can be used in
 CC.
    the diagnosis, monitoring, research, or treatment of conditions
     characterised by the expression of various cancer associated antigens.
 CC.
     The invention provides nucleic acid sequences and encoded polypeptides
 .CC
     which are cancer associated antigen precursors expressed in human breast
 CC
     cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC
     lung cancer.
 CC
 XX
     Sequence 831 BP; 285 A; 165 C; 167 G; 209 T; 5 other;
 SQ
                             Score 684.6; DB 20; Length 831;
                      67.5%;
  Query Match
                      96.1%; Pred. No. 1.9e-180;
  Best Tocal Similarity
                                              Indels
                                                          Gaps
                            0; Mismatches
                                          23;
  Matches 764; Conservative
           1. ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
             ŨУ
          37 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 96
 Db
          QΥ
          Db
         Qу
         Db
```

```
181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           QУ
       217 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 276
Db
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           Qy
       277 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 336
ηb
        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           Qу
        337 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 396
Db
        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           QУ
        297 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 456
Db
        421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           QУ
        457 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 516
Db
        481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           QУ
        517 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 576
Do
        54% TCAGATGCCTTTGCTACTTTCAA-GGATTTACTAACCAGACATAAAGTGTTGGTAGC-AG 598%
          ŮΔ
        CAGATGCCTTTGCTACTTTCAAGGGATTTACTAACCAGACATAAAGTGTTGGTAGCAAG 636
EL
        ACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 65899
            .⊜y
        ACTTCTTAGAACAAAATTACGACACTANTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 695: /
 \hat{\lambda}_{i}
        500 AG-AATTATGTTACTAAGAGACAGTCTTTAAAG-CTGCTAGGGGAGCTGATCCTGGACCG 71658
            QΥ
        597 AGAAATTATGTTACCAAGAGACAGTCCTTAAAGCCTGCTAAGGGAACTGATTCTGGACCG 756
 Db
         717 TCACAACTTTGCCATC-ATGACAAAGTATATCAGCAAGCC---GGAGAACCTGAAACTCA 772
                                             \delta \lambda
                             757 TCANAACTTTGCCATCAANGCAAAAGTTTATCAACAAGCCNGGGGAAACCGGAAACNCAA 816
 Tit:
         773 TGATGAACCTCCTTC 787
 QУ
             11 1111111111
         817 GGAGGAACCTCCTTC 831
 Db
 PESULT 7
 AAI60020/C
     AAI60020 standard; cDNA; 1191 BP.
 TD
 XX
     AAIE0020;
 AC
 XX
               (first entry)
     22-OCT-2001
 DT
 XX
     Human polynucleotide SEQ ID NO 4009.
 DE
     Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX
     peripheral nervous system; neuropathy; central nervous system; CNS;
  ΚW
  KW
```

```
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
     amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW
     chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
ΚW
KW
     leukaemia; ss.
KW
XX
     Homo sapiens.
OS
ХX
     WO200153312-A1.
PN
XX
     26-JUL-2001.
PD
XX
     26-DEC-2000; 2000WO-US34263.
PF
XX
     21-JAN-2000; 2000US-0488725.
PR
     25-APR-2000; 2000US-0552317.
PR
     09-JUL-2000; 2000US-0598042.
PR
     19-JUL-2000; 2000US-0620312.
25
     03-AUG-2000; 2000US-0653450.
PR
      14-SEP-2000; 2000US-0662191.
DΒ
      19-OCT-2000; 2000US-0693036.
 \mathbb{P} \mathbb{R}^{n}
      29-NOV-2000; 2000US-0727344.
 PR
XX
      (HYSE-) HYSEQ INC.
 PΑ
 XX
                                                               Ren F,
      Ning YE, Liu C, Asundi V. Chen R, Ma Y,
                                                    Qian XB,
 T_{\rm S}
      Many J. Wang Z, Wehrman T, Ku C, Xue AJ,
                                                     Yang Y,
                                                               Zhang J;
 TI
      Timo QA, Zhou P, Goodrich R, Drmanac PT;
 22.
      WPI; 2001-442253/47.
 157.
      9-PSDB; AAM40864.
 23.
      Movel nucleic acids and polypeptides, useful for treating disorders,
 M_{i}^{i}
 200
      queh as central nervous system injuries -
 24
 XX.
      Claim 1; SEQ ID NO 4009; 10078pp; English.
 PS
      The invention relates to human nucleic acids (AAI57798-AAI61369) and
 XΧ
      the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC
       immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC
       in gene therapy. A composition containing a polypeptide or polynucleotide
 CO
       of the invention may be used to treat diseases of the peripheral nervous
 CO
       system, such as peripheral nervous injuries, peripheral neuropathy and
 CC
       localised neuropathies and central nervous system diseases, such as
 CC
       Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC
       lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
  CC
       atilisation of the activities such as: Immune system suppression,
  da
       Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
  ÇC
       and thrombolytic activity, cancer diagnosis and therapy, drug screening,
  ĊC
       assays for receptor activity, arthritis and inflammation, leukaemias and
  CC
  CC
       Note: The sequence data for this patent did not form part of the printed
  CC
  CC
       specification.
  CC
       Sequence 1191 BP; 348 A; 261 C; 236 G; 346 T; 0 other;
  XX
  SQ:
                             67.5%; Score 684.4; DB 22; Length 1191;
    Query Match
                             99.9%; Pred. No. 2.6e-180;
    Eest Local Similarity
```

```
0;
                                                  Gaps
                                     1; Indels
                                               0:
                          Mismatches
 Matches 685; Conservative
                        0:
       329 GCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCC 388
          QУ
      1189 GCACTCGAAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCC 1130
Db
       389 TCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTA 448
          Qу
      1129 TCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTA 1070
Db
       449 TTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTA 508
          Qу
       1069 TTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTA 1010
Db
       509 AGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATT 568
           ÕΑ
       1009 AGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATT 950
Db
        569 TACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTT 628
           QV
        949 TACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTT 890
Lb
        529 TTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAA 688
           Cor
          TTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGACACAGTCTTTAA 830
Db
        AND AGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCA 74822
           27
          AGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCA 77.0
Db
        745 CCAACCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCC 308
           Qy
        759 GCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCC 710
 1 3/2
        309 AGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTA 868
           오동
        709 AGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTA 650
        869 TTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAA 928
           649 TTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAA 590
 Db
        929 AAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCC 988
            0.7
        539 AAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCC 530
 Db
         989 GAGACTTGAAGAAAACGGCCCCTTGA 1014
 QΥ
            11111111111111111111111111111
         529 GAGACTTGAAGAAAACGGCCCCTTGA 504
 Db
```

RESULT 8

AAC91772 ID AAC91772 standard; cDNA; 1026 BP.

XX

AC AAC91772;

 $\mathbb{X}\mathbb{X}$

```
27-MAR-2001 (first entry)
DT
    Human ANIC-BP (acute neuronal induced calcium-binding protein) cDNA.
ХX
DE
    Human; acute neuronal induced calcium-binding protein; ANIC-BP;
XX
     Mo25 homologue; HymA homologue; drug screening; stroke;
KW
     acute head trauma; multiple sclerosis; spinal cord injury; vaccine;
KW
KW
     cerebroprotective; neuroprotective; ss.
KW
XX
     Homo sapiens.
os
XX
     WO200078947-A1.
PN
XX
     28-DEC-2000.
PD
XX
     14-JUN-2000; 2000WO-EP05457.
PF
XX
                    99EP-0112024.
     22-JUN-1999;
ΡR
XX
      (MERE ) MERCK PATENT GMBH.
 PΑ
 XX
                                           Von Melchner L;
     Den Daas I, Fischer V, Seyfried C,
 PΙ
 XX
     WPI; 2001-102721/11.
 DR
      P-PSDB; AAB48970.
 DR.
     Novel acute neuronal induced calcium binding protein, useful for
 XX
      treating acute head trauma, stroke, multiple sclerosis and spinal
 ुनुः
 وراتي
      rjury
 3977
 XX
      Claim 5; Page 35-36; 50pp; English.
 38
      The invention relates to human acute neuronal induced calcium-binding
 XX
      protein (ANIC-BP) and to nucleic acid encoding it. The invention
 COL
      also relates to expression systems and recombinant host cells comprising
 CC
      ANIC-BP DNA, the recombinant production of ANIC-BP, antibodies specific
 CC.
      for ANIC-BP, fusion proteins comprising ANIC-BP and an immunoglobulin
 CC
      Fc region, and methods of screening for modulators of ANIC-BP function.
 CC
      ANIC-BP has homology and structural similarity to HymA and Mo25 proteins.
 CC
      ANIC-BP proteins and nucleotides are useful for treating stroke and
 CC
      acute head trauma, multiple sclerosis and spinal cord injury. ANIC-BP
 CC
      proteins are useful in screening assays, for identifying membrane bound
 CC
      or soluble receptors, and also in vaccines. ANIC-BP nucleotides are
 CC
      useful as diagnostic reagents, as tools for tissue expression studies,
  CG.
      for chromosome localisation studies, as genetic vaccines, and in
  CC
       the generation of transgenic animals. The present sequence represents
  CC
  CC
      cDNA encoding human ANIC-BP.
  CC
       Sequence 1026 BP; 359 A; 199 C; 203 G; 265 T; 0 other;
  XΧ
  SO
                                    Score 582.6; DB 22; Length 1026;
                            57.5%;
    Query Match
                                    Pred. No. 5.5e-152;
                            74.78;
    Best Local Similarity
                                                         Indels
                                      Mismatches 244;
    Matches 748; Conservative
                                   0;
             18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGC 77
                Q7
             12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 71
  Db
```

Qу	78 CATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
Db	72 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 131
ОХ	129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGTGCAAAACGAGAAAGAA
Dр	132 CAAAAATCTGGTTGCCATGAAAGAAATCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 191
Qу	189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
	AGCTGACCTGAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
ДУ	
Ολ	309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
ממ	312 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 371
Qy-	369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
Db -	
QУ ДУ	429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
e Vas	*** TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
. Sign	CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 6084
Db	
Q <u>y</u> Db	509 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
	569 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
QY F	
JL	
QУ	CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
Db	732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 791
Ğλ	789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848
ି ସପ	849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
Qу	
Db	## 100 PT

```
909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
                                                               11111111
              Qу
         912 GTTCCTCAGCAAGTTTCAGAACGACAGGACGAGAGGATGAGCAGTTTAACGACGAGAAGAC 971
Db
          969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
              Çγ
          972 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1012
Db
RESULT 9
ABK13127
     ABK13127 standard; cDNA; 3281 BP.
ID
XX
     ABK13127;
AC
XX
     09-APR-2002 (first entry)
DT
     Human secretory polynucleotide (sptm) cDNA (481257.3).
XX
DΕ
     Signal peptide; transmembrane domain; human; sptm; ss; gene;
XX
     481257.3; antiarteriosclerotic; antiatherosclerotic; antipsoriatic;
KW
      autiinflammatory; cytostatic; anti-HIV; antiallergic; antidiabetic;
 KW
      nephrotropic; antigout; antithyroid; hepatotropic; neuroprotective;
 ΚW
      osseopathic; antirheumatic; antiarthritic; dermatological; cancer;
40A
      inmunosuppressive; antiulcer; ophthalmological; vulnerary; gout;
 KW
      anticonvulsant; cerebroprotective; nootropic; antiparkinsonian;
 KM
      varucide; antibacterial; cell proliferative disorder; arteriosclerosis;
 WII
      atherosclerosis; psoriasis; immune system disorder; inflammation;
 KW
      acquired immunodeficiency syndrome; AIDS; Addison's disease;
                                                                         # C +
 VV
      adult respiratory distress syndrome; allergy; cirrhosis; osteoporosis;
 \mathbb{R}N
      diabetes mellitus; Graves' disease; multiple sclerosis; osteoarthritis;
 KW
      rheumatoid arthritis; systemic lupus erythematosus; ulcerative colitis;
 RИ
      haematopoietic cancer; neurological disorder; stroke; epilepsy;
 Huntington's disease; Parkinson's disease; meningitis; prion disease;
 KW
      kurn; Creutzfeldt-Jakob disease; cerebral palsy; myasthaenia gravis;
 3.44
 Zii
      diabetic neuropathy; Alzheimer's disease.
 KW
 YX
      Homo sapiens.
 οŻ
 XX
      WO200111032-A1.
 PN
 ΧX
      15-FEB-2001.
 PD
 XX
       01-JUN-2000; 2000WO-US15246.
  \overline{PF}
 XX
                      99US-147500P.
       05-AUG-1999;
  PR
                      99US-147501P.
       05-AUG-1999;
  PR
  XX
       (INCY-) INCYTE GENOMICS INC.
  PA
  XX
                                 Russo FD, Spiro PA,
Cohen HJ, Rosen BH,
                                                       Banville SC;
       Hodgson DM, Lincoln SE,
                                                       Chalup MS, Hillman JL;
  \mathbb{P}\mathbb{I}
       Braucher SR, Dufour GE,
       Jones AL, Yu JY, Greenawalt LB, Panzer SR,
                                                      Poseberry AM;
  PI
  PI
       Wright RJ, Daniels SE;
  PI
  XX
       WPI; 2002-147236/19.
  DR
  XX
```

Novel secretory polynucleotide (sptm) and polypeptides encoded by sptm, useful for diagnosing and treating disorders or diseases associated with cell signaling e.g., allergy, psoriasis, Grave's disease, epilepsy

Claim 1; Page 192-193; 198pp; English.

PT

PT

TT TT XX

PS

CC

CC

CC

CC

CC

CC CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Ch.:

CC

CC

CC.

CC.

 $\mathcal{E}_{\mathcal{Q}}$

Quary Match

This invention relates to novel cDNA molecules encoding isolated secretory polynucleotides (sptm) with similarity to signal peptide (SP) or transmembrane domain (TM) consensus sequences. The polynucleotide sequences of the invention are useful for producing sptm protein by recombinant techniques, the protein may be used to generate anti-sptm antibodies which may be used to analyse protein expression levels in different tissues. The sptm molecules are useful for diagnostic and therapeutic purposes e.g., to diagnose or treat a condition associated with cell signaling such as a cell proliferative disorders (e.g., arteriosclerosis, atherosclerosis, psoriasis, cancers), immune system disorders (e.g., inflammation, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, cirrhosis, diabetes mellitus, gout, Graves' disease, multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative colitis and haematopoietic cancer), a neurological disorder (e.g., stroke, epilepsy, Huntington's disease, Parkinson's disease, meningitis, prion diseases including kuru, ceutzfeldt-Jakob disease, cerebral palsy, myasthenia gravis, diabetic mouropathy and Alzheimer's disease). Sptm sequences can be used to were detect the presence of or quantifying the amount of sptm-related polynucleotide in a sample. The sptm polynucleotide is used to design probes useful in diagnostic assays carried out to detect or confirm conditions, disorders, or diseases associated with abnormal levels of . Sptm expression. Sptm, its fragments or oligonucleotides derived from optm may be used as primers in amplification steps prior to hybridisation. The present sequence represents the human sptm (481257.3) cDNA sequence of the invention.

Sequence 3281 BP; 1014 A; 601 C; 676 G; 990 T; 0 other;

57.5%;

Pred. No. 9.3e-152; 74.7%; Best Local Similarity Indels ·Gaps : 0; Mismatches 244; Matches 748; Conservative 13 GTTTAGTAAATCACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77 Ωy 101 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 160 Db 78 CATTTTGGAAAAGCAAGAC------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128 ÇΫ 161 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAAGGCTACAGAAGAAGTTTC 220 Db $G\lambda$ 221 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 280 Db 189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248 QУ]}}}| 281 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 340 Db

Score 582.6; DB 24;

Length 3281;

	QУ	249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAA 308	
OY 369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428 IIII IIII IIIIIIIIIIIIIIIIIIIIIIIIIII	_		
	-		
	- -		٠
609 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668	QУ	349: CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608	,
CACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCACACTTTGC 728-	ĞĀ	609 BCAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668	
CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788	Q y	CAS CASTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728	
ON SAT AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 940- ON SAT AGACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908- ON SAT TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908- ON SAT TCCTCACAAAACACAGCCCATCCTAGAACACTCCTCCAAAAACCAGGCCAAACTCATAGA 1000- ON SAT TCCTCACAAAAACACAGCCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1000- ON SAT TCCTCACAAAAACAAAACAAAAGAAAGAACGGCCAAACTCATAGA 1000- ON SAT TCCTCACAAAAACAAACGACCGAGATGATGAGCAGTTCGCTGACGAGAAGAA 968- ON SAT AGACAAAACCACCCCAACATTGAAAAAAAAAAAAAAA	ΩŻ		
Qy 849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908			
Oy 969 CTACTTGATTAAACAGATCCGAGACTTGAGAGAAAACGGCCC 1009	Qy,	949 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908	3 '
Oy 969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009	QУ	909 GTTTCTGAGCAGCTTCCAAAAAGAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 96	8:
	ОУ	969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009	

ID ABV22937 standard; cDNA; 3849 BP.

```
ABV22987;
AC
XX
     13-SEP-2002 (first entry)
DT
XX
     Human prostate expression marker cDNA 22978.
DΕ
XX
     Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW
KW
     pharmacogenomic marker; gene; ss.
XX
     Homo sapiens.
OS
XX
PN
     WO200160860-A2.
XX
     23-AUG-2001.
PD
XX
     20-FEB-2001; 2001WO-US05171.
PF
XX
     17-FEB-2000; 2000US-183319P.
₽R
     16-MAR-2000; 2000US-189862P.
PR
     25-MAY-2000; 2000US-207454P.
PR
     09-JUN-2000; 2000US-211314P.
₽R
     18-JUL-2000; 2000US-219007P.
PR
     13-DEC-2000; 2000US-255281P.
\Sigma R.
XX
     (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
AS
XX
     Collegel R. Endege WO,
                               Monahan JE;
ÐΤ
     WEI; 2001-662795/76.
1....
77
     Novel isolated nucleic acid molecule associated with cancerous state of
20
     prostate cells and correlating with presence of prostate cancer, useful
PT
     for detecting presence of prostate cancer, stage of prostate cancer -
\mathcal{D}^{(g)}
\Sigma\Sigma
     Claim 1; Page 4088; 11750pp; English.
MS
XX
     The invention relates to an isolated nucleic acid molecule (I) comprising
CC
     a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC
     specification or its complement. (I) is useful for:
CC
      (a) assessing whether a patient is afflicted with prostate cancer;
CC
      (b) monitoring the progression of prostate cancer in a patient;
CC
      (c) assessing the efficacy of a test compound to inhibit prostate
CC
     cancer in a patient;
CC
      (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC
CC
     in a patient;
     (e) selecting a composition for inhibiting prostate cancer in a patient;
CC
      (f) assessing the prostate cell carcinogenic potential of a compound;
CC
      (g) determining whether prostate cancer has metastasized in a patient;
CC
      (h) assessing the aggressiveness or indolence of prostate cancer in a
CC
CC
     patient;
      (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
CC
XX
     Sequence 3849 BP; 1142 A; 745 C; 858 G; 1081 T; 23 other;
30
                                   Score 582.6; DB 23; Length 3849;
                            57.5%;
  Query Match
                                   Pred. No. 1e-151;
  Best Local Similarity
                           74.7%;
                                   0; Mismatches 244; Indels
                                                                    9; Gaps
                                                                                1;
  Matches 748; Conservative
```

Qy Db		GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAACACAATTTGGC	
Qy Db		CATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTC	
Qy	129	TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	188
Db	1.89	AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT	248
Db Ωу		AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA	
Db		AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA	
Db :		CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC	
CS:		TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG	
(257) (35)	4.9	GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA	488%~
Qy Db	*	TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC	
Qy Db		CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA	
Dp		ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT	
Qy Db		TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC	
Qy Db		CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACC'ICCTTCG	
Qy Db		GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 	

to the second

```
849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
QΫ
             1277 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1336
Db
         909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
ΩУ
             1337 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1396
Db
         969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
Qу
             1861 - 11111111111 1 11 1111111 1 1 11 1
         1397 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1437
D.b
RESULT 11
ABV28822
     ABV28822 standard; cDNA; 3849 BP.
ID
XX
     ABV28822;
AC-
XX
     16-SEP-2002 (first entry)
DT
XX -
     Human prostate expression marker cDNA 28813.
DE
XX
     Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
\mathbb{Z}M
     pharmacogenomic marker; gene; ss.
W
ZX
     Homo sapiens.
េន
XX
     WO200160860-A2.
\mathbb{P}N
\mathbb{I}X
     23-AUG-2001.
     20-FEB-2001; 2001WO-US05171.
25
XX
     17-FEB-2000; 2000US-183319P.
92
     16-MAR-2000; 2000US-189862P.
55
     25-MAY-2000; 2000US-207454P.
₽R
     09-JUN-2000; 2000US-211314P.
 PR
     13-JUL-2000; 2000US-219007P.
 ĿR
     13-DEC-2000: 2000US-255281P.
 PR
 XX
      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA
 ХX
     Schlegel R, Endege WO,
                              Monahan JE;
 PT.
 XX
     WPI; 2001-662795/76.
 DR
 XX
     Novel isolated nucleic acid molecule associated with cancerous state of
 ΤÇ
      prostate cells and correlating with presence of prostate cancer, useful
 PT
      for detecting presence of prostate cancer, stage of prostate cancer -
 PT
 XX
      Claim 1; Page 6066-6067; 11750pp; English.
 PS
 XX
      The invention relates to an isolated nucleic acid molecule (I) comprising
 CC
      a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC
      specification or its complement. (I) is useful for:
 CC
      (a) assessing whether a patient is afflicted with prostate cancer;
 CC
```

```
(b) monitoring the progression of prostate cancer in a patient;
CC
    (c) assessing the efficacy of a test compound to inhibit prostate
CC
    cancer in a patient;
CC
    (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC
    in a patient;
CC
    (e) selecting a composition for inhibiting prostate cancer in a patient;
CC
    (f) assessing the prostate cell carcinogenic potential of a compound;
CC
    (g) determining whether prostate cancer has metastasized in a patient;
CC
    (h) assessing the aggressiveness or indolence of prostate cancer in a
CC
    patient;
CC
    (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
CC
XX
    Sequence 3849 BP; 1142 A; 745 C; 858 G; 1081 T; 23 other;
SQ
                           Score 582.6; DB 23; Length 3849;
                     57.5%;
 Query Match
                           Pred. No. 1e-151;
                     74.78;
 Best Local Similarity
                                                       Gaps
                                                              1;
                          0; Mismatches 244;
                                            Indels
 Matches 748; Conservative
         18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
QΥ
           437 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 496
ďŒ
                                  AAAAAGACAGACAAGGCTTCAGAAGAAG'IGTC 128
         78 CATTTTGGAAAAGCAAGAC
ЭŻ
                                  TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 556
٦b
        CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 616
DU
        AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
OY
                                                     AGAAGCAGTAGCTCAACTTCCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 676.
ಾಗಿ
        308.
QУ
            677 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 736
Db.
        368 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC
QУ
               737 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 796
Db.
        369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
QÝ
                  1111 177 1 1 1 11111 11111 1 1111 1 111 1
        797 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 856
Do
        429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
QУ
                             1 1
             857 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 916
Db
        439 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
 QУ
                    917 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 976
 Db
         549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
 QY
             11111 | 1 | 1 | 1 |
         977 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 1036
 Db
```

```
609 ACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
                             QУ
           111 1111 11 1 1111
       1037 ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCAGAAAATTATGT 1096
Db
        669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
            Qy
       Db
        729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
            QУ
       1157 PATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1216
Db
        789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848
            QY
       1217 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTTGTAGCCAA 1276
Db
        849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
            Ωy
        1277 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1336
Db
        909 GTTTCTGAGCAGCTTCCAAAAAGAAAGAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
            QУ
        1337 GTTCCTCAGCAAGTTTCAGAACGACAGGACGAGAGGAGTGAGCAGTTTAACGACGAGAAGAC 1396
 dG
         D69 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
            \bigcirc
        1397 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1437
 D.
 RESULT 12
     AAF30688 standard; cDNA; 1053 BP.
 ED
 XX
     AME30688; . .
 AC
 XY
     11-JUN-2001 (first entry)
 DT.
     Human acute neuronal induced calcium binding protein ANIC-BP-1B cDNA.
 -31-11
-23-11
 125
     Acute neuronal induced calcium binding protein; ANIC-BP-1B;
 XX
     spice variant; human; stroke; head trauma; Parkinson's disease;
 KII
     Alzheimer's disease; multiple sclerosis; spinal cord injury;
 KN
     cerebroprotective; antiparkinsonian; nootropic; neuroprotective;
 KM
 KW
     therapy; diagnosis; vaccine; ss.
 КW
 X:
     Homo sapiens.
 OS
 XX
                  Location/Qualifiers
      Key-
 FH
                   1..1053
      CDS
  FT
                   /*tag=a
  FT
                   /product= "Human ANIC-BP-1B"
  FT
  XX
      WO200125423-A1.
  ΡN
  XX
      12-APR-2001.
  РĐ
  XX
```

```
28-SEP-2000; 2000WO-EP09475.
PF
XX
                  99EP-0119113.
    04-OCT-1999;
PR
XX
    (MERE ) MERCK PATENT GMBH.
PA
XX
    Duecker K, Den Daas I;
ΡI
XX
    WPI; 2001-266306/27.
DR
    P-PSDB: AAB20387.
DR
XX
    Novel human acute neuronal induced calcium-binding protein like protein
PT
    splice variant, useful for treating stroke, acute head trauma,
PT
    Parkinson's disease, Alzheimer's disease multiple sclerosis, spinal
PT
    cord injury -
pT
XX
    Claim 4; Page 43-44; 49pp; English.
PS
XX
    The present sequence is that of cDNA encoding a novel human acute
CC
    neuronal induced calcium binding protein-like protein splice
ec.
    variant, ANIC-NP-1B (see AAB20387). The protein shows homology to
CC
    other members of the calcium binding protein family, including
CC
    PMIC-BP, a protein discovered by mRNA differential display that is
CC
    upregulated in a rat model of head trauma. ANIC-BP and ANIC-BP-1B
CC
    differ in their C-terminal portions. The variant protein could
CC
    Merve as a novel drug target. The invention provides ANIC-BP-1B
CK!
     polynucleotides and polypeptides, expression vectors, host cells
CC.
     antibodies, as well as methods for producing the protein and
10.1
     for treating or preventing disorders associated with expression of to
CC
     the protein by inhibiting or activating the action of ANIC-BP-1B. . . . .
₽Ç
     seases that may be treated include stroke and acute head trauma, D. .
CC
     warkinson's disease, Alzheimer's disease, multiple sclerosis and
CC.
     pinal cord injury. The polynucleotides and polypeptides can also ex
100
     be used in diagnostic assays and in vaccines, and to identify
ČÜ
     agonists and antagonists useful for treating conditions associated
Or.
     with ANIC-BP-1B imbalance.
ďC
XX
     Sequence 1053 BP; 357 A; 211 C; 214 G; 271 T; 0 other;
S\Omega
                        53.4%; Score 541.6; DB 22; Length 1053;
  Query Match
                        74.1%; Pred. No. 1.5e-140;
  Best Local Similarity
                              0; Mismatches 239; Indels
                                                           11.; Gaps
  Matches 716; Conservative
          18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
 QУ
             12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGCATGGC 71
 ĐΩ
                                       AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
           78 CATTTTGGAAAAGCAAGAC--
 Ωу
                                       11 1111111111
           72 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 131
 Do
          QV
                         132 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 191
 Db
          189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
 QУ
```

To be	192 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 251	
Db	THE ACCEPTANCE CONTRACT AND CONTRACT ACCEPTANCE A	
Qy.	2/19 AGCTGACCTGCAGCTGATAGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	
Dp	252 AGCTGATTTACAGCTCATTGACTTTTAGGGGGGTTTAGTGCTCATCC 368 300 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368	
Qy		
Dр		
Qу	369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428	
Db		
QУ	429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488	
Db		
QУ	489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548	
Db		
ON.	549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608 	
DM ·		
ű y	ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCIGAGAATTATGT 669	
වර්		
⊘y	TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728	
7 ² 23		
<u>O</u> y	729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 78	
D.b	732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGGT	
СХ	789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 84	
Db	792 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTGTTG	
Q ₂ /	849 TCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 90	
Гр	852 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 91	
Qy	909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAG 96	
Ep		
QУ	967 AACTAC 972	
Dh		•

```
AAS89557 standard; cDNA; 1162 BP.
ID
XX
AC
     AAS89557;
XX
     13-FEB-2002
                   (first entry)
DT
XX
     DNA encoding novel human diagnostic protein #25361.
DE
XX
     Human; chromosome mapping; gene mapping; gene therapy; forensic;
ΚW
     food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW
XX
OS
     Homo sapiens.
XX
     WO200175067-A2.
ΡN
XX.
PD
     11-GCT-2001.
XX
     30-MAR-2001; 2001WO-US08631.
PF
XX
     31-MAR-2000; 2000US-0540217.
₽R
     23-AUG-2000; 2000US-0649167.
PR
XX
     (HYSE-) HYSEQ INC.
\mathbb{P}\mathbf{A}
XX
                           Tang YT;
     Drmanac RT, Liu C,
PΤ
XX
                                                                      7 5
     WPI; 2001-639362/73.
\partial \mathbb{R}
     P-PSDB, ABG25370.
્
XX
     New isolated polynucleotide and encoded polypeptides, useful in
\mathbf{PT}
     diagnostics, forensics, gene mapping, identification of mutations
22
     responsible for genetic disorders or other traits and to assess
ЭŤ
     biodiversity
2.12
MK.
     Claim 1; SEQ ID No 25361; 103pp; English.
pg.
XX
     The invention relates to isolated polynucleotide (I) and
CC
     polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC
     polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC
     and gene mapping, and in recombinant production of (II). The
CC
     polynucleotides are also used in diagnostics as expressed sequence tags
CC
     for identifying expressed genes. (I) is useful in gene therapy techniques
CC
     to restore normal activity of (II) or to treat disease states involving
CC
      (II). (II) is useful for generating antibodies against it, detecting or
CC
     quantitating a polypeptide in tissue, as molecular weight markers and as
CC
     a food supplement. (II) and its binding partners are useful in medical
CC
      imaging of sites expressing (II). (I) and (II) are useful for treating
CC
     disorders involving aberrant protein expression or biological activity.
CC
     The polypeptide and polynucleotide sequences have applications in
CC
     diagnostics, forensics, gene mapping, identification of mutations
CC
      responsible for genetic disorders or other traits to assess biodiversity
CC
      and to produce other types of data and products dependent on DNA and
CC
      amino acid sequences. AAS64197-AAS94564 represent novel human
cc.
      diagnostic coding sequences of the invention.
CC
     Note: The sequence data for this patent did not appear in the printed
CC
      specification, but was obtained in electronic format directly from WIPO
```

at ftp.wipo.int/pub/published_pct_sequences.

CC

CC

Sequence 1162 BP; 383 A; 241 C; 258 G; 280 T; 0 other;

	Query Ma Best Loc Matches	in for	milarity Conservat	73.9%;	Pred.	539.6; No. 5.6e matches	-140;			Gaps	4;
) (2)		1	TTTAGTAAATC	- 11111	11111			1 1 1 1	111	1111	
) (10)			ATTTTGGAAAA 	111111			1111			11111	
Ď.			TAAATCACTGCA	11 11	11111	11111	111 111	11 11 11			
Q D			AGAAGCAGTGGG 	1111 11	11 111	111111	i 1	11 11 11	1 1 1		,
(1)	Z D	249 1	· · COCTCN CCTCC	AGCTGATÀ	GACTTT	GAGGGAAA 	AAAAGAT	GTGACCCA	GATATI	TAACAA	308
Ç	Y	34/9	CATCTTGAGAA CATTCTCAGAA	GACAGATA	AGGCA-C	TCGGAGTC 	CTACTGT	GGAGTATA	ATTAGTO	ECTCATC	3670
Ç		T68	CTCA TATCC	TGTTTAT(CTCCTC	AAAGGATA	TGAAGCC	CCACAGAT	TGCC	CTTACGT	423
ζ	n Ny Oir	₹ 3 ,4	TGTGGGATTAT TGTGGAATAA1	GCTGAGA(GAATGTA	TTCGACAT	GAACCA(CTTGCCAA	AATCAT(CCTCTTT	483
ζ)y Ob	484	TCTAATCAA'TT TCGGAACAG'TT	'CAGAGAT'	TTCTTTA	AAGTACGTO	GAGTTG	rcaacatt'	TGATAT	rgcttca 	543
Ç	Ωγ Ob	544	GATGCCTTTGC	TACTTTC	AAGGAT'I	TTACTAAC(CAGACA I	AAAGTGTT	GGTAGC	AGACTTC	603
(Qy Db	604	TTAGAACAAAA TTGGAACAGCA	ATTACGAC	ACTATT	TTTGAAGA(CTATGAG	AAATTGCT	TCAGTC	TGAGAAT	663
	Qy Db		TATGTTACTA	1 1 1 1 1				11		11111	
	ĞΆ		TTTGCCATCA'						'CATGAT		

```
863 TTCACAATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTG 922
\Gamma b
         784 CTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTG 843 .
Qу
             923 CTGCGAGACAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTA 982
         844 GCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTC 903
Qy
             983 GCCAATCCTAACAAGACGCCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTC 1042
         904 ATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAG 963
QУ
             1043 ATAGAGTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAG 1102
         964 AAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
Ċν
             1103 AAGACCTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1148
Db
RESULT 14
AAX39818/c
    AAX39818 standard; DNA; 833 BP.
ID:
XX
ÄС
     AAX39818;
XX
     02-JUL-1999 (first entry)
LT
\Sigma \Sigma
     Castric cancer associated gene.
Sin
     Cancer associated antigen; diagnosis; research; treatment; human;
KW.
     breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KN
     prostate cancer; ss.
Y_{ij}^{h}
XX
OS.
     Memo sapiens.
XX
     %09904265-A2.
ЪИ
X\Sigma
     20-JAN-1999.
DD
\Sigma X
                   98WO-US14679.
     15-JUL-1998;
 PF
XX.
     22-JUN-1998; 98US-0102322.
 \mathbf{p}_{\mathrm{R}}
                   97US-0896164.
 ŔŖ
     17-JUL-1997;
                   97US-0061599.
ÐÑ.
     10-OCT-1997;
     10-OCT-1997;
                   97US-0061765.
 \mathbb{PR}
                   97US-0948705.
     10-OCT-1997;
 טת
                   97GB-0021697.
     11-OCT-1997;
 PP.
 XX
     (LUDW-) LUDWIG INST CANCER RES.
 PΑ
 XX
     Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
 ÐΙ
     Pfrewidschuh M, Sahin U, Scanlan MJ, Stockert E;
 PI
     Tureci O;
 рΤ
 XX
     WPI; 1999-132448/11.
 DŘ
 XX
     New isolated cancer associated nucleic acids and polypeptides -
 PT
```

isolated using sera from cancer patients, used to develop products PT for the diagnosis, monitoring or treatment of cancers PT XXClaim 67; Page 559; 787pp; English. PS The invention relates to a method for diagnosing a disorder characterised XXby expression of a human cancer associated antigen precursor coded for by CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a CC biological sample isolated from a subject with an agent that specifically CC CC binds to the NAM, an expression product or a fragment of an expression product complexed with an HLA molecule; and (b) determining the CC CC interaction between the agent and the NAM or the expression product as a CC determination of the disorder. The products and methods can be used in CC the diagnosis, monitoring, research, or treatment of conditions characterised by the expression of various cancer associated antigens. CCCC The invention provides nucleic acid sequences and encoded polypeptides which are cancer associated antigen precursors expressed in human breast CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and CC CClung cancer. CC XXSequence 833 BP; 253 A; 171 C; 172 G; 227 T; 10 other; SO Length 833; Score 520.2; DB 20; 51.3%; Query Match Pred. No. 1.2e-134; 98.1%; Best Local Similarity Gaps Indels Mismatches 7: Natches 566; Conservative 0; 442 GAATGTATTCGACATGAACCACTTG-CCAAAATCATCCTC-TTTTCTAATCAAT,TCAGAG 499 QW. 732 JAATNTATTCGACTTGACCCANTTGCCCAAANTCATCCTCTTTTTCTAATCAATTCAGAG 673 500 ATTTCTTTAAGT-ACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558 ATTTCTTTAAGTAACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 613 Db 559 TTCAAGGATTTACTAACCAGA-CATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTA 617 QΨ 612 TTCAAGGATTTACTAACCNGACCTTAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTA 553 Db 618 CGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAG 677 Q7 552 CGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAG 493 UD 678 ACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGAC 737 Øу £92 ACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGAC 433 nh 738 AAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAG 797 Qy 432 AAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAG 373 D.b 798 TCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAA 857 Ôλ 372 TCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAA 313 Db 858 AACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAG 917

312 AACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAG 253

Qу

Db

```
918 CAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGAT 977
              QУ
         252 CAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGAT 193
Db
          978 TAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              Q7
          192 TAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 156
Ν'n
RESULT 15
AAS88031
     AAS88031 standard; cDNA; 2492 BP.
ID
XX
     AAS88031;
AC
XX
                  (first entry)
     13-FEB-2002
DT
XX
     DNA encoding novel human diagnostic protein #23835.
DE
     Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX
     food supplement; medical imaging; diagnostic; genetic disorder; ss.
 WX
 ΧW
 XX
     Romo sapiens.
 OS
 XX
      %21.00175067-A2.
 \mathbb{P}N
 XY
      2:-OCT-2001.
 DD
 ....- НАК-2001; 2001WO-US08631.
 82
 تأييده
      1-MAR-2000; 2000US-0540217.
 Ŭ3-AUG-2000: 2000US-0649167.
 TL
 \Sigma X
      (HYSE-) HYSEQ INC.
 PΑ
 XX
      Drmanac RT, Liu C,
                           Tang YT;
 PΙ
 XX
      WPI; 2001-639362/73.
 DR
      P-PSDB; ABG23844.
 DR
      New isolated polynucleotide and encoded polypeptides, useful in
  XΧ
      diagnostics, forensics, gene mapping, identification of mutations
  ЭT
      responsible for genetic disorders or other traits and to assess
  ψŢ
  TG
      biodiversity
  PT
  XX
       Claim 1; SEQ ID No 23835; 103pp; English.
  ୁଦ୍ର
       The invention relates to isolated polynucleotide (I) and
  XX
       polypeptide (II) sequences. (I) is useful as hybridisation probes,
  CC
       polymerase chain reaction (PCR) primers, oligomers, and for chromosome
  CC
       and gene mapping, and in recombinant production of (II). The
  CC
       polynucleotides are also used in diagnostics as expressed sequence tags
  CC
       for identifying expressed genes. (I) is useful in gene therapy techniques
  CC
       to restore normal activity of (II) or to treat disease states involving
  CC
       (II). (II) is useful for generating antibodies against it, detecting or
  CC
       quantitating a polypeptide in tissue, as molecular weight markers and as
  CC
```

CC

```
a food supplement. (II) and its binding partners are useful in medical
    imaging of sites expressing (II). (I) and (II) are useful for treating
CC
    disorders involving aberrant protein expression or biological activity.
CC
    The polypeptide and polynucleotide sequences have applications in
CC
    diagnostics, forensics, gene mapping, identification of mutations
CC
    responsible for genetic disorders or other traits to assess biodiversity
CC
    and to produce other types of data and products dependent on DNA and
CC
    amino acid sequences. AAS64197-AAS94564 represent novel human
CC
    diagnostic coding sequences of the invention.
CC
    Note: The sequence data for this patent did not appear in the printed
CC
    specification, but was obtained in electronic format directly from WIPO
CC
    at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
    Sequence 2492 BP; 751 A; 477 C; 546 G; 718 T; 0 other;
SQ
                     48.9%; Score 496; DB 23;
                                           Length 2492;
 Ouery Match
                           Pred. No. 1.1e-127;
 Best Local Similarity
                     73.0%;
                                                               7:
                           0; Mismatches 255;
                                            Indels
                                                    18;
                                                        Gaps.
 Matches 737; Conservative
         18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
Qу
                            143 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 202
Db
                                 --AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
         78 CATTTTGGAAAAGCAAGAC--
Q7
                                  11 11111111111
        203 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 262
Db
        ÚΖ
                     CAAAAATCTGGTTGCCATGAAAGAATTCTGTATGGCACAAATGAAAAAGATCGTCAGAC 3226
DD
        AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
QY
            323 AGAAGCAGCAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTATCACCCTGGT 382
Db
        249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 30.8
Qy.
            383 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 442
Db
        309 CATCTTGAGAAGACAGATAGGCA-CTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATC 367
Q7
               43 TATTCTCAGAAGACAAATTGGTACCGAGAACTCCTACTGTTGAATACATCTGCACCCAAA 502
Db
         358 CTCA--TATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATT--GCCTTACGT 423
ÇΫ
                     1 1
                                                       11 1
         503 CAGAATATTTTTGTTCATGTTATTGAAAGGGTATGAATCTCCCAGAAATAGCTCTAAATT 562
Db
         424 TGTGGGATTATGCTGAGAAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTT 483
Qy
            563 TGTGGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGGCAAAATCATTTTGTGG 622
Db
         434 TCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCA 543
 ÕА
                        623 TCGGAACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCA 682
 Db
         544 GATGCC-TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTT 602
 QY.
```

CC

)b	683	GATGCCATTTGCCACATTCAAGGGTTTACTTACAAGACATAAATTGCTCAGTGCAGTATT	
Qγ	603	CTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAA	662
Dh	743		802 [.]
Qy	663	TTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAA	722
Db	803		862
Qy	723	CTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db	863		922
Qy		CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGG-TGTT	
ďQ		CTGCTGCGAGACAAAGTCGCAACATCCAGTTTTGAGGCCTTTCACGTTTTTAAGGCAGTT	
δÃ. · · ·		TGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA	
Db		TGTAGCCAATCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCTAT	
Ç <u>v</u>		ACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGA	
Db		ACTCATAGAGTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTTTTTT	1102
Ćλ		CGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009	
Db	1.103	CGAGAAGACCTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1152	

Search completed: January 6, 2004, 01:28:41

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 01:16:12; Search time 94 Seconds

(without alignments)

4761.303 Million cell updates/sec

Title: US-10-088-872-1

Perfect score: 1014

Sequence: 1 atgaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:*

1: /cgn2 6/ptodata/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2 6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2 6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2 6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

			8				
Res	ult		Query				
	No.	Score	Match	Length	DB	ID	Description
	1	1014	100.0	1421	4	US-09-620-312D-111	Sequence 111, App
	2	1010.8	99.7	1344	3	US-09-190-965-2	Sequence 2, Appli
	3	1010.8	99.7	1344	4	US-09-470-253-2	Sequence 2, Appli
С	4	51.6	5.1	7218	1	US-08-232-463-14	Sequence 14, Appl
	5	37.8	3.7	1457	4	US-09-214-307A-9	Sequence 9, Appli
С	6	37.4	3.7	4103	4	US-09-620-312D-390	Sequence 390, App
	7	36.4	3.6	4533	3	US-08-726-214-5	Sequence 5, Appli
С	8	35.6	3.5	577	4	US-09-513-057C-20	Sequence 20, Appl
С	9	35.6	3.5	3707	3	US-09-276-531-42	Sequence 42, Appl
С	10	35.6	3.5	5714	4	US-09-620-312D-393	Sequence 393, App
	11	35.2	3.5	12793	4	US-09-004-838-124	Sequence 124, App

	12	35	3.5	775	3	US-08-961-083-89	Sequence 89, Appl
	13	35	3.5	775	4	US-09-536-784-89	Sequence 89, Appl
	14	35	3.5	1696	3	US-08-961-083-217	Sequence 217, App
	15	35	3.5	1696	4	US-09-536-784-217	Sequence 217, App
	16	35	3.5	3763	1	US-07-792-865D-1	Sequence 1, Appli
С	17	35	3.5	13926	4	US-08-961-527-5	Sequence 5, Appli
С	18	34.8	3.4	1413	4	US-09-016-434-1382	Sequence 1382, Ap
	19	34.8	3.4	12734	4	US-09-344-456-1	Sequence 1, Appli
	20	34.8	3.4	14078	3	US-09-433-262-1	Sequence 1, Appli
	21	34.8	3.4	14078	4	US-09-702-330-1	Sequence 1, Appli
	22	34.8	3.4	14578	3	US-08-859-694-1	Sequence 1, Appli
С	23	34.6	3.4	64467	4	US-09-803-671B-3	Sequence 3, Appli
	24	34.4	3.4	22481	4	US-08-367-841A-43	Sequence 43, Appl
	25	34.4	3.4	22481	5	PCT-US95-07201-43	Sequence 43, Appl
	26	34.4	3.4	22484	4	US-09-875-223-2	Sequence 2, Appli
	27	34	3.4	3378	4	US-09-328-352-4107	Sequence 4107, Ap
С	28	34	3.4	43360	4	US-09-453-702B-206	Sequence 206, App
С	29	34	3.4	45325	4	US-09-453-702B-261	Sequence 261, App
	30	33.8	3.3	189	4	US-09-134-001C-1145	Sequence 1145, Ap
С	31	33.8	3.3	4233	3	US-09-056-105-27	Sequence 27, Appl
	32	33.6	3.3	1276	3.	US-09-177-325-2	Sequence 2, Appli
	33	33.6	3.3	1276	3	US-09-411-812A-2	Sequence 2, Appli
	34	33.6	3.3	1276	4	US-09-590-113-2	Sequence 2, Appli
	35	33.6	3.3	1347	4	US-09-134-001C-1309	Sequence 1309, Ap
	36	33.4	3.3	1410	4	US-09-328-352-124	Sequence 124, App
С	37	33.4	3.3	1508	1	US-08-236-311-5	Sequence 5, Appli
C	38	33.4	3.3	1508	3	US-08-457-918-5	Sequence 5, Appli
	39	33.4	3.3	2643	4	US-09-486-072-6	Sequence 6, Appli
С	40	33.4	3.3	18627	4	US-08-961-527-113	Sequence 113, App
	41	33.2	3.3	378	4	US-09-252-991A-10313	Sequence 10313, A
С	42	33.2	3.3	813	4	US-09-252-991A-10547	Sequence 10547, A
	43	33.2	3.3	831	4	US-09-252-991A-10512	Sequence 10512, A
С	44	33.2	3.3	3253	4	US-09-333-214-4	Sequence 4, Appli
	45	33.2	3.3	8091	4	US-09-230-652-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-620-312D-111

- ; Sequence 111, Application US/09620312D
- ; Patent No. 6569662
- ; GENERAL INFORMATION:
- ; APPLICANT: Tang, Y. Tom
- ; APPLICANT: Liu, Chenghua
- ; APPLICANT: Asundi, Vinod
- APPLICANT: Zhang, Jie
- ; APPLICANT: Ren, Feiyan
- : APPLICANT: Chen, Rui-hong
- ; APPLICANT: Zhao, Qing A.
- ; APPLICANT: Wehrman, Tom
- ; APPLICANT: Xue, Aidong J.
- ; APPLICANT: Yang, Yonghong
- ; APPLICANT: Wang, Jian-Rui
- ; APPLICANT: Zhou, Ping
- ; APPLICANT: Ma, Yunqing

```
APPLICANT:
          Wang, Dunrui
          Wang, Zhiwei
  APPLICANT:
          John Tillinghast
  APPLICANT:
          Drmanac, Radoje T.
  APPLICANT:
  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
  TITLE OF INVENTION: Polypeptides
  FILE REFERENCE: 784CIP2B
  CURRENT APPLICATION NUMBER: US/09/620,312D
  CURRENT FILING DATE: 2000-07-19
  PRIOR APPLICATION NUMBER: 09/552,317
  PRIOR FILING DATE: 2000-04-25
  PRIOR APPLICATION NUMBER: 09/488,725
  PRIOR FILING DATE: 2000-01-21
  NUMBER OF SEQ ID NOS: 1105
  SOFTWARE: pt FL genes Version 1.0
 SEO ID NO 111
  LENGTH: 1421
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (217)..(1230)
US-09-620-312D-111
 Query Match
                  100.0%; Score 1014; DB 4;
                                      Length 1421;
 Best Local Similarity
                  100.0%;
                        Pred. No. 4.8e-292;
 Matches 1014; Conservative
                       0: Mismatches
                                   0;
                                                      0;
                                      Indels
                                                Gaps
        1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Qу
         217 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 276
Db
       Qу
         Db
       Qу
         Db
       181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qу
         397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
Db
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Qу
         Db
       457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516
       301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Qу
         Db
       517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
       361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Qу
         577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
Db
       421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
Qy
```

Db	637	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	696
Qу	481	${\tt TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT}$	540
Db	697		756
Qу	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600
Db	757		816
Qу	601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	660
Db	817		876
Qу	661	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	720
Db	877		936
Qу	721	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db	937		996
Qу	781	$\tt CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT$	840
Db	997	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	1056
QУ	841	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	900
Db	1057	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	1116
Qy ,	901	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	960
Db	1117	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	1176
QУ	961	GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014	
Db	1177	GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1230	

US-09-190-965-2

- ; Sequence 2, Application US/09190965
- ; Patent No. 6071721
- ; GENERAL INFORMATION:
- ; APPLICANT: Tang, Y. Tom
- ; APPLICANT: Guegler, Karl J.
- ; APPLICANT: Corley, Neil C.
- ; APPLICANT: Gorgone, Gina A.
- : TITLE OF INVENTION: CALCIUM BINDING PROTEIN
- ; FILE REFERENCE: PF-0635 US
- ; CURRENT APPLICATION NUMBER: US/09/190,965
- ; CURRENT FILING DATE: 1998-11-13
- ; NUMBER OF SEQ ID NOS: 5
- ; SOFTWARE: PERL Program
- ; SEQ ID NO 2
- ; LENGTH: 1344

```
TYPE: DNA
;
  ORGANISM: Homo sapiens
  FEATURE: -
  OTHER INFORMATION: 3734805
US-09-190-965-2
                 99.7%; Score 1010.8; DB 3; Length 1344;
 Query Match
                 99.8%; Pred. No. 4.2e-291;
 Best Local Similarity
                                           0; Gaps
                                                   0;
                                 2:
 Matches 1012; Conservative
                     0; Mismatches
                                    Indels
        1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
QУ
         124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
Db
       Qу
         Db
      Qу
         244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAAGAA 303
Db
      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qу
         304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363
Db
      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Qу
         364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423
Db
      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Qу
         424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483
Db
      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Qу
         484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543
Db
      421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
Qу
         544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603
Db
      481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
Qv
         604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663
Db
      541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
Qу
         664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723
Db
      601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
Qу
         724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783
Db
      661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Qу
         784 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843
```

Db

```
721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
Qy
          844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903
Db
       781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Qу
          904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963
Db
       841 GTGGCCAGTCCTCACAAAACACGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
QУ
           964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023
Dh
       901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Qy
          1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083
Db
       961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QУ
           1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
Db
RESULT 3
US-09-470-253-2
; Sequence 2, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
  APPLICANT: Tang, Y. Tom
  APPLICANT: Guegler, Karl J.
  APPLICANT:
           Corley, Neil C.
  APPLICANT: Gorgone, Gina A.
  TITLE OF INVENTION: CALCIUM BINDING PROTEIN
  FILE REFERENCE: PF-0635 US
  CURRENT APPLICATION NUMBER: US/09/470,253
  CURRENT FILING DATE: 1999-12-22
  PRIOR APPLICATION NUMBER: 09/190,965
  PRIOR FILING DATE: 1998-11-13
  NUMBER OF SEQ ID NOS: 5
  SOFTWARE: PERL Program
 SEQ ID NO 2
   LENGTH: 1344
   TYPE: DNA
   ORGANISM: Homo sapiens
   FEATURE: -
   OTHER INFORMATION: 3734805
US-09-470-253-2
                    99.7%; Score 1010.8; DB 4; Length 1344;
 Query Match
 Best Local Similarity 99.8%; Pred. No. 4.2e-291;
                                                           0;
 Matches 1012; Conservative
                         0; Mismatches
                                       2;
                                          Indels
                                                  0; Gaps
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
QУ
           124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
Db
        Qу
```

Db	184	CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	243
Ωу	121	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGTG	180
Db	244		303
QУ	181	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG	240
Db	304	CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG	363
Qу	241	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	300
Db	364	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	423
Qy	301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db	424	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	483
Qу	361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
Db	484	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	543
ДУ	421	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	480
Db	544	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	603
ДÀ	481	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db .	604		663
Qу	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600
Db	664		723
ΣУ	601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	660
Db	724	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	783
ΩУ	661	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	720
Db	784	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	843
ДÀ	721	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db	844	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	903
Ο̈́У	781	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840
Db	904	$\tt CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTT$	963
Qу	841	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	900
Db	964	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	1023
Qу	901	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	960
Db	1024	$\tt CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC$	1083

Best Local Similarity 3.6%; Pred. No. 4.2e-05;

Query Match

5.1%; Score 51.6; DB 1; Length 7218;

```
Matches 12; Conservative 196; Mismatches 130; Indels 0; Gaps
                                              0;
       1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Qy
        Db
      Qу
         Db
      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAAGAA 180
Qу
        Db
     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qу
           Db
      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Qу
        Db
     301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAG 338
QУ
          Db
RESULT 5
US-09-214-307A-9
; Sequence 9, Application US/09214307A
; Patent No. 6544516
; GENERAL INFORMATION:
 APPLICANT: NEUTEC PHARMA PLC
  TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF INFECTIONS OF GRAM POSITIVE
  TITLE OF INVENTION: COCCI
  FILE REFERENCE: PM 259204
  CURRENT APPLICATION NUMBER: US/09/214,307A
  CURRENT FILING DATE: 1999-01-04
  PRIOR APPLICATION NUMBER: PCT/GB97/01830
; PRIOR FILING DATE: 1997-07-07
; PRIOR APPLICATION NUMBER: GB9614274.0
  PRIOR FILING DATE: 1996-07-06
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
  LENGTH: 1457
  TYPE: DNA
  ORGANISM: Staphylococcus aureus
US-09-214-307A-9
                3.7%; Score 37.8; DB 4; Length 1457;
 Query Match
 Best Local Similarity 47.0%; Pred. No. 0.23;
 Matches 150; Conservative 0; Mismatches 167; Indels 2; Gaps
      430 ATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
Qу
         386 AATATGAGAACTGTAGTTGATCGACCTAGAACACAATATAAAAAAGTCGTCTTTAATAAT 445
Db
```

```
490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
Qу
              446 TTATTTTATCAATTTAGTAAGGATGCCAACTTTGAACCTATTGCTTGTAGACCCTATCGT 505
Db
        550 TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA 609
Qу
              506 CCTCAAACAAAGGGTCTGTTGAATCATTAGCTAAATTTGTTGAACAGCGTTTAAGACCA 565
Db
        610 CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT 669
Qy
             566 TACGATTATGAATTTTATGATGCTG--TAGAACTTATTGGGCTAGTAAACGATTTATGTC 623
Db
        670 ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCC 729
Qу
                      Db
        624 ACGAATTGAATCACTTAGAAATTTCACAAGCAACAGAACAACGACCTATCGACGTTTTCA 683
Qу
        730 ATCATGACAAAGTATATCA 748
            Db
        684 ATTATGAAGAAAAAGAACA 702
RESULT 6
US-09-620-312D-390/c
; Sequence 390, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
  APPLICANT: Liu, Chenghua
  APPLICANT: Asundi, Vinod
  APPLICANT: Zhang, Jie
  APPLICANT: Ren, Feiyan
  APPLICANT: Chen, Rui-hong
  APPLICANT: Zhao, Qing A.
  APPLICANT: Wehrman, Tom
  APPLICANT: Xue, Aidong J.
  APPLICANT: Yang, Yonghong
  APPLICANT: Wang, Jian-Rui
  APPLICANT: Zhou, Ping
  APPLICANT: Ma, Yunging
  APPLICANT: Wang, Dunrui
  APPLICANT: Wang, Zhiwei
  APPLICANT: John Tillinghast
  APPLICANT: Drmanac, Radoje T.
  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
  TITLE OF INVENTION: Polypeptides
  FILE REFERENCE: 784CIP2B
  CURRENT APPLICATION NUMBER: US/09/620,312D
  CURRENT FILING DATE: 2000-07-19
  PRIOR APPLICATION NUMBER: 09/552,317
  PRIOR FILING DATE: 2000-04-25
  PRIOR APPLICATION NUMBER: 09/488,725
  PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
 SOFTWARE: pt FL genes Version 1.0
; SEO ID NO 390
   LENGTH: 4103
```

```
TYPE: DNA
   ORGANISM: Homo sapiens
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (104)..(3493)
US-09-620-312D-390
                       3.7%; Score 37.4; DB 4; Length 4103;
 Query Match
 Best Local Similarity 60.2%; Pred. No. 0.53;
         62; Conservative 0; Mismatches 41; Indels 0; Gaps
                                                                   0;
 Matches
          6 AAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAA 65
Qу
                      | | |
        4091 ACAAATGAGAAAGTTTCATTTACCTCAAAAAAATCCAGGCTATACAAACAGACAACTGAA 4032
Db
         Qу
            4031 AGCCACATAGGAAATTTCCGAAACACAAAAGAAAAAGTCTCAC 3989
Db
RESULT 7
US-08-726-214-5
; Sequence 5, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
    APPLICANT: Tang, Wei-Jen
    APPLICANT: Gilman, Alfred G.
ï
    TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;
    TITLE OF INVENTION: AND USES THEREFOR
;
    NUMBER OF SEQUENCES: 31
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Arnold, White & Durkee
     STREET: P.O. Box 4433
     CITY: Houston
     STATE: Texas
     COUNTRY: United States of America
;
      ZIP: 77210
;
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/726,214
      FILING DATE: Concurrently Herewith
      CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 60/005,498
      FILING DATE: 04-OCT-1995
    ATTORNEY/AGENT INFORMATION:
      NAME: Highlander, Steven L.
      REGISTRATION NUMBER: 37,642
      REFERENCE/DOCKET NUMBER: UTSD: 450
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (512) 418-3000
      TELEFAX: (512) 474-7577
```

INFORMATION FOR SEQ ID NO: 5:

```
SEQUENCE CHARACTERISTICS:
     LENGTH: 4533 base pairs
;
     TYPE: nucleic acid
;
     STRANDEDNESS: single
     TOPOLOGY: linear
US-08-726-214-5
 Query Match
                      3.6%; Score 36.4; DB 3; Length 4533;
 Best Local Similarity 56.8%; Pred. No. 1.1;
 Matches 67; Conservative 0; Mismatches 51; Indels
                                                     0; Gaps
                                                                 0;
        718 CACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATG 777
Qу
                   1 11 1
       2644 CTCATCGCCACCATCATGCTGGTGCAGGTCAGCCACATGGTGAAGCTGACACTCATGCTG 2703
Db
        778 AACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGG 835
Qу
                   Db
       2704 CTCGTCACAGGCGCCGTGACTGCCATCAACCTGTATGCCTGGTGTCCTGTCTTTGATG 2761
RESULT 8
US-09-513-057C-20/c
; Sequence 20, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND
PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
  CURRENT APPLICATION NUMBER: US/09/513,057C
  CURRENT FILING DATE: 2000-02-24
  NUMBER OF SEQ ID NOS: 35
 SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
  LENGTH: 577
   TYPE: DNA
   ORGANISM: Lycopersicon esculentum
US-09-513-057C-20
                      3.5%; Score 35.6; DB 4; Length 577;
 Query Match
 Best Local Similarity 51.2%; Pred. No. 0.64;
        83; Conservative 0; Mismatches 79; Indels 0; Gaps
                                                                 0;
        457 GAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTG 516
Qу
            223 GACCCAAATACCCAAAACACATCTTTACATAGAAATCAAGAGATTTCTGAAGCACACAG 164
Db
        517 GAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACC 576
Qу
                  163 AAGCAAAAAGATGTATAATTTCACAAAATTACTATTATATTTTTCTGTGATCATGTAAC 104
Db
        577 AGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
QУ
            11 1 1
                   103 AGGCCTTGTTGGTAAGCACAATAATATGAAGAAAGAGATTAC 62
Db
```

```
US-09-276-531-42/c
; Sequence 42, Application US/09276531
; Patent No. 6183968
  GENERAL INFORMATION:
    APPLICANT: Bandman, Olga
    APPLICANT: Lal, Preeti
    APPLICANT: Hillman, Jennifer L.
    APPLICANT: Yue, Henry
    APPLICANT: Reddy, Roopa
    APPLICANT: Guegler, Karl J.
     APPLICANT: Baughn, Mariah R.
     TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
     TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL
PROLIFERATION
    NUMBER OF SEQUENCES: 134
     CORRESPONDENCE ADDRESS:
;
       ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;
       STREET: 3174 PORTER DRIVE
      CITY: PALO ALTO
       STATE: CALIFORNIA
       COUNTRY: USA
       ZIP: 94304
     COMPUTER READABLE FORM:
       MEDIUM TYPE: Floppy disk
       COMPUTER: IBM PC compatible
;
       OPERATING SYSTEM: PC-DOS/MS-DOS
;
       SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;
     CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/276,531
      FILING DATE: Herewith
      CLASSIFICATION:
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER: 60/079,677
       FILING DATE: March 27, 1998
;
       CLASSIFICATION:
;
     ATTORNEY/AGENT INFORMATION:
       NAME: Lynn E. Murry, Ph.D.
       REGISTRATION NUMBER: 42,918
       REFERENCE/DOCKET NUMBER: PA-0008 US
     TELECOMMUNICATION INFORMATION:
       TELEPHONE: (650) 855-0555
       TELEFAX: (650) 845-4166
   INFORMATION FOR SEQ ID NO: 42:
     SEQUENCE CHARACTERISTICS:
       LENGTH: 3707 base pairs
;
       TYPE: nucleic acid
       STRANDEDNESS: single
       TOPOLOGY: linear
     IMMEDIATE SOURCE:
       LIBRARY: CERVNOT01
       CLONE: 936117
US-09-276-531-42
                          3.5%; Score 35.6; DB 3; Length 3707;
  Query Match
  Best Local Similarity 51.9%; Pred. No. 1.7;
                                                                             0;
           80; Conservative 0; Mismatches
                                                 74; Indels
                                                                 0: Gaps
```

```
Qу
        445 TGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTC 504
             Db
        3154 TGCTTTCAAAATGTGGAACAAACTAAAATATAAGGCTTTTCTGATAAACTATAAAAATTT 3095
        505 TTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAG 564
Qу
               3094 AATCAGCACTTGGATCTAATGACATATCTTTATAATACTTCCTCTGCAGATACATTCACT 3035
Db
         565 GATTTACTAACCAGACATAAAGTGTTGGTAGCAG 598
Qy
              Db
        3034 TAGTTCAAACCTTAACATACAAAGTTAGTCTCAG 3001
RESULT 10
US-09-620-312D-393/c
; Sequence 393, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
  APPLICANT: Tang, Y. Tom
  APPLICANT: Liu, Chenghua
  APPLICANT: Asundi, Vinod
  APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
;
  APPLICANT: Chen, Rui-hong
;
  APPLICANT: Zhao, Qing A.
  APPLICANT: Wehrman, Tom
  APPLICANT: Xue, Aidong J.
  APPLICANT: Yang, Yonghong
  APPLICANT: Wang, Jian-Rui
  APPLICANT: Zhou, Ping
  APPLICANT: Ma, Yunqing
  APPLICANT: Wang, Dunrui
  APPLICANT: Wang, Zhiwei
  APPLICANT: John Tillinghast
  APPLICANT: Drmanac, Radoje T.
  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
  TITLE OF INVENTION: Polypeptides
  FILE REFERENCE: 784CIP2B
  CURRENT APPLICATION NUMBER: US/09/620,312D
  CURRENT FILING DATE: 2000-07-19
  PRIOR APPLICATION NUMBER: 09/552,317
  PRIOR FILING DATE: 2000-04-25
  PRIOR APPLICATION NUMBER: 09/488,725
  PRIOR FILING DATE: 2000-01-21
  NUMBER OF SEQ ID NOS: 1105
  SOFTWARE: pt FL genes Version 1.0
; SEQ ID NO 393
   LENGTH: 5714
   TYPE: DNA
   ORGANISM: Homo sapiens
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (272)..(4312)
US-09-620-312D-393
 Query Match
                       3.5%; Score 35.6; DB 4; Length 5714;
 Best Local Similarity 51.9%; Pred. No. 2.2;
```

```
Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps
                                                                      0:
        445 TGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTC 504
Qу
            5233 TGCTTTCAAAATGTGGAACAAACTAAAATATAAGGCTTTTCTGATAAACTATAAAAATTT 5174
Db
        505 TTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAG 564
Qу
              5173 AATCAGCACTTGGATCTAATGACATATCTTTGTAATACTTCCTCTGCAGATACATTCACT 5114
Db
        565 GATTTACTAACCAGACATAAAGTGTTGGTAGCAG 598
Qy
             5113 TAGTTCAAACCTTAACATACAAAGTTAGTCTCAG 5080
Db
RESULT 11
US-09-004-838-124
; Sequence 124, Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
    APPLICANT: Michelmore, Richard W.
    APPLICANT: Shen, Kathy
;
    APPLICANT: Meyers, Blake
    TITLE OF INVENTION: Procedures and Materials for
    TITLE OF INVENTION: Conferring Pest Resistance in Plants
    NUMBER OF SEQUENCES: 140
    CORRESPONDENCE ADDRESS:
;
      ADDRESSEE: Townsend and Townsend and Crew LLP
;
      STREET: Two Embarcadero Center, Eighth Floor
      CITY: San Francisco
      STATE: California
     COUNTRY: USA
      ZIP: 94111-3834
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
;
      OPERATING SYSTEM: PC-DOS/MS-DOS
;
      SOFTWARE: PatentIn Release #1.0, Version #1.30
;
    CURRENT APPLICATION DATA:
;
    APPLICATION NUMBER: US/09/004,838
      FILING DATE: 09-JAN-1998
      CLASSIFICATION: 800
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/781,734
      FILING DATE: 10-JAN-1997
    ATTORNEY/AGENT INFORMATION:
      NAME: Einhorn, Gregory P.
      REGISTRATION NUMBER: 38,440
      REFERENCE/DOCKET NUMBER: 023070-078810US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (415) 576-0200
      TELEFAX: (415) 576-0300
   INFORMATION FOR SEQ ID NO: 124:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 12793 base pairs
;
      TYPE: nucleic acid
;
      STRANDEDNESS: single
```

```
TOPOLOGY: linear
;
    MOLECULE TYPE: DNA
    FEATURE:
      NAME/KEY: -
      LOCATION: 1..12793
      OTHER INFORMATION: /note= "RG2S"
US-09-004-838-124
 Query Match
                      3.5%; Score 35.2; DB 4; Length 12793;
 Best Local Similarity 47.6%; Pred. No. 4.4;
 Matches 101; Conservative 10; Mismatches
                                           98; Indels
        438 GAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAG 497
Qy
            Db
       5998 GAGARAGWAWGRRRGAKAKARMCSMSYTTGGGATGTGATACTTCTTTTAGGAAAATGGAG 6057
        498 AGATTTCTTTAAGTACGTGGAGTTGTCA---ACATTTGATATTGCTTCAGATGCCTTTGC 554
Qy
              Db
       6058 TTATATCTTTGATATTGTATTTTTTAATGTAATTTATATATTTTAATCATTTTAGTTTAT 6117
Qу
        555 TACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAA 614
             6118 AAGTTTTATTTATTTTGATATGAAAAAAAAAGTCTTTTATACATTGGATTTAACATAAAA 6177
Db
        615 TTACGACACTATTTTTGAAGACTATGAGAAAT 646
Qу
             Db
       6178 ATCCAACAATATTAATCAAAAAGACCAMACAT 6209
RESULT 12
US-08-961-083-89
; Sequence 89, Application US/08961083
; Patent No. 6159469
  GENERAL INFORMATION:
    APPLICANT: Choi et. al.
    TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
    NUMBER OF SEQUENCES: 452
    CORRESPONDENCE ADDRESS:
     ADDRESSEE: Human Genome Sciences, Inc.
     STREET: 9410 Key West Avenue
     CITY: Rockville
     STATE: Maryland
     COUNTRY: USA
     ZIP: 20850
;
    COMPUTER READABLE FORM:
     MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
     COMPUTER: HP Vectra 486/33
     OPERATING SYSTEM: MSDOS version 6.2
     SOFTWARE: ASCII Text
    CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/961,083
     FILING DATE:
     CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
    APPLICATION NUMBER:
     FILING DATE:
    ATTORNEY/AGENT INFORMATION:
```

```
NAME: Brookes, A. Anders
;
     REGISTRATION NUMBER: 36,373
     REFERENCE/DOCKET NUMBER: PB340P2
    TELECOMMUNICATION INFORMATION:
     TELEPHONE: (301) 309-8504
     TELEFAX: (301) 309-8512
  INFORMATION FOR SEQ ID NO: 89:
    SEQUENCE CHARACTERISTICS:
     LENGTH: 775 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: double
     TOPOLOGY: linear
US-08-961-083-89
                      3.5%; Score 35; DB 3; Length 775;
 Query Match
 Best Local Similarity 46.9%; Pred. No. 1.1;
                                                       2; Gaps
 Matches 143; Conservative 0; Mismatches 160; Indels
         22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGCCATT 81
Qу
            263 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 322
Db
         82 TTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
Qу
                   1 11
        323 GCGCCAAGAGAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 382
Db
        142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCCAACAGAAGCAGTGGCT 201
Qу
            383 GAAGAGAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 442
Db
        202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
Qу
                443 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 500
Db
        262 CTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
QУ
             501 TTGAAACGCCTGCTGTAGAAAAACAAACAAACAAACCAACAGAGGAACCAAAAGTTGAACAAG 560
Db
        322 CAGAT 326
QУ
            \square
        561 CAGGT 565
Db
RESULT 13
US-09-536-784-89
; Sequence 89, Application US/09536784
; Patent No. 6573082
   GENERAL INFORMATION:
       APPLICANT: Choi et. al.
       TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
       NUMBER OF SEQUENCES: 452
       CORRESPONDENCE ADDRESS:
            ADDRESSEE: Human Genome Sciences, Inc.
            STREET: 9410 Key West Avenue
            CITY: Rockville
            STATE: Maryland
            COUNTRY: USA
```

```
ZIP: 20850
;
       COMPUTER READABLE FORM:
           MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
           COMPUTER: HP Vectra 486/33
           OPERATING SYSTEM: MSDOS version 6.2
           SOFTWARE: ASCII Text
       CURRENT APPLICATION DATA:
           APPLICATION NUMBER: US/09/536,784
            FILING DATE: 30-Oct-1997
           CLASSIFICATION: <Unknown>
       PRIOR APPLICATION DATA:
           APPLICATION NUMBER: 08/961,083
            FILING DATE: OCT-30-1997
       ATTORNEY/AGENT INFORMATION:
           NAME: Michelle S. Marks
            REGISTRATION NUMBER: 41,971
            REFERENCE/DOCKET NUMBER: PB340P3
       TELECOMMUNICATION INFORMATION:
           TELEPHONE: (301) 309-8504
           TELEFAX: (301) 309-8512
   INFORMATION FOR SEQ ID NO: 89:
     SEQUENCE CHARACTERISTICS:
           LENGTH: 775 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: double
            TOPOLOGY: linear
       SEQUENCE DESCRIPTION: SEQ ID NO: 89:
US-09-536-784-89
                      3.5%; Score 35; DB 4; Length 775;
 Query Match
 Best Local Similarity 46.9%; Pred. No. 1.1;
 Matches 143; Conservative 0; Mismatches 160; Indels
                                                       2; Gaps
         22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
Qу
            263 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 322
Db
         82 TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
Qγ
             323 GCGCCAAGAGAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 382
Db
        142 GCAATGAAAGAATTCTGTGTGGTACAAACGAGAAGAACCCCCAACAGAAGCAGTGGCT 201
Qу
            383 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 442
Db
        202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
QУ
            443 GAAGAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 500
Db
        262 CTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
Qу
             501 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 560
Db
        322 CAGAT 326
Qу
            11111
Db
        561 CAGGT 565
```

```
RESULT 14
US-08-961-083-217
; Sequence 217, Application US/08961083
; Patent No. 6159469
  GENERAL INFORMATION:
    APPLICANT: Choi et. al.
    TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
    NUMBER OF SEQUENCES: 452
;
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Human Genome Sciences, Inc.
      STREET: 9410 Key West Avenue
      CITY: Rockville
      STATE: Maryland
      COUNTRY: USA
      ZIP: 20850
    COMPUTER READABLE FORM:
;
      MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;
      COMPUTER: HP Vectra 486/33
;
      OPERATING SYSTEM: MSDOS version 6.2
      SOFTWARE: ASCII Text
    CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/961,083
      FILING DATE:
      CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
     APPLICATION NUMBER:
      FILING DATE:
    ATTORNEY/AGENT INFORMATION:
      NAME: Brookes, A. Anders
      REGISTRATION NUMBER: 36,373
      REFERENCE/DOCKET NUMBER: PB340P2
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (301) 309-8504
      TELEFAX: (301) 309-8512
;
  INFORMATION FOR SEQ ID NO: 217:
    SEQUENCE CHARACTERISTICS:
;
      LENGTH: 1696 base pairs
;
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
US-08-961-083-217
                        3.5%; Score 35; DB 3; Length 1696;
 Ouerv Match
  Best Local Similarity 46.9%; Pred. No. 1.7;
 Matches 143; Conservative 0; Mismatches 160; Indels
                                                          2; Gaps
                                                                  1;
         22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGCCATT 81
Qy
                                         275 AGTCAACCATCAGACAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 334
Db
         82 TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
Qу
                    335 GCGCCAAGAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 394
Db
        142 GCAATGAAAGAATTCTGTGTGGTACAAACGAGAAGAACCCCCAACAGAAGCAGTGGCT 201
Qу
```

```
Db
         395 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 454
         202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
Qу
                  Db
         455 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 512
         262 CTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
Qу
              Db
         513 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 572
         322 CAGAT 326
Qу
             111 1
Db
         573 CAGGT 577
RESULT 15
US-09-536-784-217
; Sequence 217, Application US/09536784
; Patent No. 6573082
   GENERAL INFORMATION:
        APPLICANT: Choi et. al.
        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
        NUMBER OF SEQUENCES: 452
        CORRESPONDENCE ADDRESS:
             ADDRESSEE: Human Genome Sciences, Inc.
             STREET: 9410 Key West Avenue
             CITY: Rockville
             STATE: Maryland
             COUNTRY: USA
             ZIP: 20850
        COMPUTER READABLE FORM:
             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
             COMPUTER: HP Vectra 486/33
             OPERATING SYSTEM: MSDOS version 6.2
             SOFTWARE: ASCII Text
        CURRENT APPLICATION DATA:
             APPLICATION NUMBER: US/09/536,784
             FILING DATE: 30-Oct-1997
             CLASSIFICATION: <Unknown>
        PRIOR APPLICATION DATA:
             APPLICATION NUMBER: 08/961,083
             FILING DATE: OCT-30-1997
        ATTORNEY/AGENT INFORMATION:
             NAME: Michelle S. Marks
             REGISTRATION NUMBER: 41,971
             REFERENCE/DOCKET NUMBER: PB340P3
        TELECOMMUNICATION INFORMATION:
             TELEPHONE: (301) 309-8504
             TELEFAX: (301) 309-8512
   INFORMATION FOR SEQ ID NO: 217:
        SEQUENCE CHARACTERISTICS:
             LENGTH: 1696 base pairs
             TYPE: nucleic acid
             STRANDEDNESS: double
             TOPOLOGY: linear
        SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-09-536-784-217
```

	cal S	3.5%; Score 35; DB 4; Length 1696; Similarity 46.9%; Pred. No. 1.7; 3; Conservative 0; Mismatches 160; Indels 2; Gaps	1;
Qу	22	AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT	81
Db	275	AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC	334
QУ	82	TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA	141
Db	335		394
QУ	142	GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCCAACAGAAGCAGTGGCT	201
Db	395		454
ΟУ	202	CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG	261
Db	455		512
QУ	262	CTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA	321 .
Db	513	TTGAAACGCCTGCTGTAGAAAACAAACAGACCAACAGAGGAACCAAAAGTTGAACAAG	572
Qу	322	CAGAT 326	
Db	573	CAGGT 577	

Search completed: January 6, 2004, 03:19:48 Job time: 96 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on:

January 6, 2004, 02:35:04; Search time 1394 Seconds

(without alignments)

2517.743 Million cell updates/sec

Title:

US-10-088-872-1

Perfect score:

1014

Secuence:

1 atgaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched:

2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters:

4526836

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database .

Published_Applications_NA:*

- /cqn2 6/ptcdata/1/pubpna/US07_PUBCOMB.seq:*
- /cgn2 6/ptodata/1/pubpna/PCT NEW PUB.seq:*
- /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:* 4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:* 5:
- /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:* 7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:* /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:* 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

옿 Query

No.

Score Match Length DB

Description

US-10-117-722-111 Sequence 111, App 1 1014 100.0 1421 13 Sequence 111, App US-10-037-270-111 15 1014 100.0 1421 2 Sequence 2, Appli US-10-025-730-2 1344 15 99.7 3 1010.8 US-09-918-995-5343 Sequence 5343, Ap 475 398 39.3 11 4 US-09-910-943-318 Sequence 318, App 28.5 690 9 5 238.8 Sequence 5263, Ap US-09-867-701-5263 24.3 435 10 246.4 6 Sequence 5899, Ap 447 10 US-09-867-701-5899 7 244.8 24.1 Sequence 4953, Ap 450 10 US-09-867-701-4953 8 244.8 24.1 Sequence 35, Appl 9 US-09-910-943-35 20.8 762 9 210.8 Sequence 1453, Ap 387 10 US-09-954-456-1453 19.2 195 10 С Sequence 481, App 387 10 US-09-880-107-481 195 19.2 11 Sequence 118, App 722 1.3 US-10-257-826A-118 169.8 16.7 1.2 US-10-257-826A-119 Sequence 119, App 700 13 166.6 16.4 13 Sequence 592, App 15.4 861 9 US-09-770-445-592 156 14 Sequence 1251, Ap 7.3 262 9 US-09-923-876-1251 74.2 15 Sequence 1251, Ap US-09-923-876-1251 7.3 262 12 74.2 16 Sequence 19069, A 11 US-09-918-995-19069 6.5 336 65.6 17 US-10-242-355-323 Sequence 323, App 6.4 187 12 -1.865 Sequence 54, Appl US-10-080-254-54 65 5.4 487 13 ï.9 Sequence 13369, A 254 10 US-09-878-574-13369 5.3 20 53.6 Sequence 777, App US-09-770-961-777 486 11 5.0 21 50.6 Ċ Sequence 2528, Ap. US-09-923-876-2528 9 . 41.6 4.1 242 22 Sequence 2528, Ap US-09-923-876-2528 41.6 4.1 242 12 23 Sequence 294, App US-10-310-154-294 1295 12 40.8 4.0 24 Sequence 1007, Ap 4.0 113306 US-10-292-798-1007 1.2 25 40.2Sequence 5787, Ap-.US-09-918-995-5787 431 11 ... 6 39.8 3.9 US-10-311-455-26 Sequence 26, Appl: 6301 2.3 1.7 11.4 3.8 Sequence 261235,13 US-10-027-632-261235 1200 13 ΓS 55.2 3.8 Sequence 261235, US-10-027-632-261235 1200 14 29 39.2 3.8 Sequence 9, Appli US-10-054-968-9 1457 15 30 37.8 3.7 Sequence 278, App US-09-873-367C-278 7178 13 01 37.8 3.7 Sequence 45, Appl US-10-001-843-45 3.7 1267 14 32 37.5 Sequence 1, Appli 3.7 3673778 13 US-10-312-841-1 33 37.6 Sequence 45, Appl US-10-087-464-45 2232 1.5 3.7 34 37.4 Sequence 335, App US-09-876-889-335 4012 9 35 37.4 3.7 C US-10-117-722-390 Sequence 390, App 4103 13 3.6 37.4 3.7 C Sequence 390, App US-10-037-270-390 4103 15 37 37.4 3.7 \mathbb{C} Sequence 1760, Ap US-10-311-455-1760 8577 13 37.4 3.7 38 Ç. Sequence 261978,. US-10-027-632-261978 13 3.7 869 39 3.7.2 €. Sequence 261978, US-10-027-632-261978 14 37.2 3.7 869 C 40 Sequence 538, App US-10-311-455-538 3.6 5413 13 41 37 $^{\rm C}$ Sequence 944, App US-10-311-455-944 3.6 9367 3.3 42 36.6 Sequence 143, App 14 US-10-079-623-143 3.6 461 43. 36.4 Sequence 29299, A 2641 12 US-10-369-493-29299 3.5 7.4. 36.4 Sequence 297, App 6071. 13 US-10-311-455-297 3.6 45 36.4

ALIGNMENTS

RESULT 1

US-10-117-722-111

[;] Sequence 111, Application US/10117722

[;] Publication No. US20030219744A1

[;] GENERAL INFORMATION:

[;] APPLICANT: Tang, Y. Tom

```
APPLICANT: Liu, Chenghua
 APPLICANT:
          Asundi, Vinod
          Zhanq, Jie
 APPLICANT:
          Drmanac, Radoje T.
 APPLICANT:
 TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 784CIP2BCIP
 CURRENT APPLICATION NUMBER: US/10/117,722
 CURRENT FILING DATE: 2002-04-04
 PRIOR APPLICATION NUMBER: 09/620,312
 PRIOR FILING DATE: 2000-07-19
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1104
 SOFTWARE: pt_FL_genes Version 1.0
 SEQ ID NO 111
  LENGTH: 1421
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (217) .. (1230)
US-10-117-722-111
                                       Length 1421;
                  100.0%;
                        Score 1014;
                                 DB 1.3;
 Cuery Match
                        Pred. No. 1.2e-281;
                  100.0%;
 Rest Local Similarity
                       0; Mismatches
                                                Gaps ...
                                    0;.
                                      Indels
 Matches 4014; Conservative
        1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
QY
          217 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 27.6
Db
       Qy
          7)
       ŢУ
          D.b
       131 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qy
          397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
∂b
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300.
Qy
          457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516
Db
       301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
QУ
          517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
Dρ
       361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
QУ
          577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
Db
```

```
421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
         Qy
       637 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 696
Db
       481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
          QУ
       697 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 756
Dp
       541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
          Q7
       757 FCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 816
Db
       501 TTCTTAGAACAAAA'ITACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
          ÖΛ
       317 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 876
Dρ
       661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
          Qу
       877 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC
Do
       721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
          QУ
          AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 996
Db
          CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840]
          W
          CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 10567
Db
       GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 9000
          1057 GTGGCCAGTCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1116
       301 CTCATTGAGT*ITCTGAGCAGCTTCCAAAAAGAAAGGACGGATGA'IGAGCAGTTCGCTGAC 960
          ्र
       1117 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1176
DJ,
       961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QУ
          1177 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1230
Ωb
```

US-10-037-270-111

- ; Sequence 111, Application US/10037270
- ; Publication No. US20030104529A1
- ; GENERAL INFORMATION:
- ; APPLICANT: Tang, Y. Tom
- APPLICANT: Liu, Chenghua
- . APPLICANT: Asundi, Vinod
- ; APPLICANT: Zhang, Jie
- ; APPLICANT: Ren, Feiyan
- APPLICANT: Chen, Rui-hong
- ; APPLICANT: Zhao, Qing A.
- : APPLICANT: Wehrman, Tom
- ; APPLICANT: Xue, Aidong J.
- ; APPLICANT: Yang, Yonghong

```
Wang, Jian-Rui
 APPLICANT:
          Zhou, Ping
 APPLICANT:
 APPLICANT:
          Ma, Yunqing
          Wang, Dunrui
 APPLICANT:
          Wang, Zhiwei
 APPLICANT:
          Tillinghast, John
 APPLICANT:
 APPLICANT: Drmanac, Radoje T.
 TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 784CIP2B
 CURRENT APPLICATION NUMBER: US/10/037,270
 CURRENT FILING DATE: 2002-01-04
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1104
 SOFTWARE: pt_FL_genes Version 1.0
 SEC ID NO 111
  LENGTH: 1421
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/HEY: CDS
  LOCATION: (217) .. (1230)
08-10-037-270-111 +
                                      Length 1421;
                        Score 1014; DB 15;
                  100.08;
 Ouery Match
                 100.0%;
                        Pred. No. 1.2e-281;
 Rest Local Similarity
                      0; Mismatches
                                               Gaps
                                                     0.;
                                   0;
                                      Indels
 Maliches 1014; Conservative
        ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
         ATGANAAAAATCCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 276
٩
       ij'n.
         Eb
       QУ
          Do
       131 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qv
          CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
25
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300.
Qу
          457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516
Db
       301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Çiy.
          517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
Db
       361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
QΥ
```

)b		577	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCCACAGATTGCCTTA	636
ðλ			CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	
DÞ:		637	CGTTGTGGGATTATGCTGAGAAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	696
Qу			TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGGAGTTGTCAACATTTGATATTGCT	
Db		697	TTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	756
ДУ			TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	
Db		757	TCAGATGCCTTTGCTACTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	816
QY			TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	
Db		217	TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	876
Qу			AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	
Db		877	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	936
Qγ			AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	
do			AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCIGAAACICAIGMIGME	
Qу	٠.	781 957	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	
			GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 	
) 			CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	
Oy Db			CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGCTTCCAAAAAGAAAG	
ΟУ			GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAÁGAAAACGGCCCCTTGA 1014	
Db		117		

US-10-025-730-2

- ; Sequence 2, Application US/10025730
- ; Publication No. US20030045466A1
- ; GENERAL INFORMATION:
- APPLICANT: Tang, Y. Tom
- ; APPLICANT: Guegler, Karl J.
- APPLICANT: Corley, Neil C.
- ; APPLICANT: Gorgone, Gina A.
 - TITLE OF INVENTION: CALCIUM BINDING PROTEIN
- ; FILE REFERENCE: PF-0635 US
 - CURRENT APPLICATION NUMBER: US/10/025,730
- CURRENT FILING DATE: 2001-12-18
- ; PRIOR APPLICATION NUMBER: US/09/190,965

```
PRIOR FILING DATE: 1998-11-13
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: PERL Program
 SEQ ID NO 2
  LENGTH: 1344
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE: -
  OTHER INFORMATION: 3734805
US-10-025-730-2
                     Score 1010.8; DB 15; Length 1344;
                99.7%;
 Query Match
                     Pred. No. 1e-280;
                99.2%;
 Best Local Similarity
                                                 0:
                                           Gaps
                       Mismatches
                                   Indels
 Matches 1012; Conservative
       1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Ωy
         124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
Db
      QY
         Do
      OV
         cC.
      ±81 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qy
         194 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363
Db
      ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
         ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423
Cb
      201 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
QУ
         424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483
      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
ÕА
         #84 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543
Db
      421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
⊋у
         564 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603
Db
      481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
Qу
         504 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663
D_{\mathcal{O}}
      541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
QУ
         664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723
Db
      601 TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
Qy
```

```
724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783
Db
        661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Qу
            734 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843
Db
        721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
Qу
            844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903
Db
        781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Qу
            904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963
Db
        841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Qv
            964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023
Db
        901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
QУ
            1024 CICATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCIGAC 1083
Db
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
ÇУ
            1984 CAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
Db
ARSULT 4
TUS-09-012-995-5343
Sequence 343, Application US/09918995
  Pablication No. US20030073623A1
                                                           1761. 1
 COMERAL INFORMATION:
  APPLICANT: Hyseq, Inc.
   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
  TITLE OF INVENTION: FROM VARIOUS CONA LIBRARIES
   FILE REFERENCE: 20411-756
   CURRENT APPLICATION NUMBER: US/09/918,995
   CURRENT FILING DATE: 2001-07-30
   PRIOR APPLICATION NUMBER: US/09/235,076
   PRIOR FILING DATE: 1999-01-20
   NUMBER OF SEQ ID NOS: 38054
   SOFTWARE: FastSEQ for Windows Version 3.0
  SEC ID NO 5343
    LENGTH: 475
    TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (1)...(475)
    OTHER INFORMATION: n = A, T, C or G
US-09-918-995-5343
                      39.3%; Score 398; DB 11; Length 475;
  Query Match
                      100.0%; Pred. No. 3.5e-104;
  Best Local Similarity
                            0; Mismatches
                                            0;
                                               Indels
                                                           Gaps
  Matches 398; Conservative
         617 ACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGA 676
```

QУ

·ij. .

```
1 ACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGA 60
Dh
       677 GACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGA 736
Qу
           61 GACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGA 120
Db
        737 CAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAA 796
Qу
           121 CAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAA 180
Db
        797 GTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACA 856
Qу
           181 GTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACA 240
Db
        857 AAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGA 916
Qy
           241 AAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGA 300
Db
        917 CCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGA 976
ÓУ
           301 GCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGA 360
Db
        977 TTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Ωy
           351 TTAAACAGATCCCAGACTTGAAGAAAACGGCCCCTTGA 398
235
RESULT 5
US-09-910-243-318
 Quence 318, Application US/09910943
                                                       ta...
  ##stent No. US20020081610A1
                                                       1544
  EXECUTED INFORMATION:
   APPLICANT: Hemmati-Brivanlou, Ali
  **PLICANT: Altman, Curtis .
  THILE OF INVENTION: Assays and Materials for Embryonic Gene Expression
  FILE REFERENCE: 7529/1G148US1
   CURRENT APPLICATION NUMBER: US/09/910,943
   CURRENT FILING DATE: 2001-07-23
   NUMBER OF SEQ ID NOS: 742
   SOFTWARE: PatentIn version 3.1
 SEC ID NO 318
   LENGTH: 690
   TYPE: DNA
    ORGANISM: Xenopus laevis
    FEATURE:
    NAME/KEY: misc feature
    LOCATION: (1)..(690)
    OTHER INFORMATION: n may be a or g or c or t/u
 US-09-910-943-318
                     28.5%; Score 288.8; DB 9;
                                            Length 690;
  Query Match
  Best Local Similarity 80.5%; Pred. No. 1.4e-72;
                                                              0;
                           0; Mismatches 82; Indels
                                                        Gaps
 Matches 338; Conservative
         595 GCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAG 654
 QУ
```

```
69 GCAGAATTTCTAGAGCAAAATTACGACAGAATATTTAATGACTATGAAAAGCTTCTTCAC 128
Db
        655 TCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGAC 714
           Qу
        129 TCTGAGAACTATGTGACGAAGAGACAGTCCCTTAAGCTGCTGGGCGAGCTGATCCTGGAC 188
Db
        715 CGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATG 774
           Qу
        189 CGACACACTTTTCCATTATGACTAAATACATAAGCAAGCCTGAAAATCTGAAGCTCATG 248
Db
        775 ATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAG 834
           Qу
        249 ATGAATCTGCTCCGTGATAAGAGCCCAAACATTCAGTTTGAAGCATTCCATGTGTTTAAG 308
Db
        835 GTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAG 894
           Qу
        309 GTGTTTGTAGCAAATCCAAACAAAACACAGCCCATCGTGGATATCCTGTTAAAAAACCAA 368
Db
        .QY
        369 ACCAAGTTAATCGACTTCCTGAGCAGCTTTCAGAAGGATCGAACAGATGACGAACAGTTC 428
Dp
        955 GC'IGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
            QУ
        *30 ACCGACGAGAAGAACTACTTGATCAAACAGATACGAGACTTAAAAAAGCCCACGCCATGA 488
Db
RESULT
ປຣະນຸ9-887 -701-5263
  (equence 5263, Application US/09867701
                                                       100
 / Patent No. US20020132237A1
                                                       SEN.
  PERAL INFORMATION:
   APPLICANT: Aglate, Paul A.
   APPLICANT: Jones, Robert
   APPLICANT: Harlocker, Susan L.
   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
   TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
   FILE REFERENCE: 210121.497
   CURRENT APPLICATION NUMBER: US/09/867,701
   CURRENT FILING DATE: 2001-05-29
   NUMBER OF SEQ ID NOS: 10912
   SOFTWARE: FastSEQ for Windows Version 4.0
  SEO ID NO 5263
    LENGTH: 435
    TYPE: DNA
    ORGANISM: Homo sapien
 US-09-867-701-5263
                           Score 246.4; DB 10; Length 435;
                      24.3%;
   Query Match
                     77.6%; Pred. No. 1.8e-60;
  Best Local Similarity
                                             Indels
                           0; Mismatches 86;
  Matches 298; Conservative
         626 TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
                  QV
          41 TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC 100
 ďŒ
         686 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
 QУ
```

```
101 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATCACAAAATACA 160
Db
       746 TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
QУ
           161 TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 220
מת
        806 TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGC 865
ΩУ
           221 TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 280
D'n
        866 CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
Qy
           231 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 340
nb
        926 AAAAAGAAAGCACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
           341 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 400
Dh
        986 TCCCAGACTTGAAGAAAACGGCCC 1009
           401 TCCGGGATTTGAAGAGACCCGCTC 424
Db
RESOLT 7
08-09-867-701-5899
/ Sequence 2009, Application US/09867701
 Patent No. - #$20020132237A1
 CONERAL INFORMATION:
  PLICANT: Aglate, Paul A.
  PPLICANT: Jones, Robert
  Harlocker, Susan L.
  MINTLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
  WITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
PINE REFERENCE: 210121.497
  CURRENT APPLICATION NUMBER: US/09/867,701
  CURRENT FILING DATE: 2001-05-29
  NUMBER OF SEQ ID NOS: 10912
  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5899
   LENGTH: 447
   TYPE: ONA
   ORGANISM: Homo sapien
US-09-367-701-5899
                    24.1%; Score 244.8; DB 10; Length 447;
 Query Match
  Best Local Similarity 77.3%; Pred. No. 5.3e-60;
  Matches 297; Conservative 0; Mismatches 87; Indels
                                                   0; Gaps
        626 TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
QΥ
           41 TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC 100
DD.
        686 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
Qу
            101 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATGACAAAATACA 160
рb
```

```
746 TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
Qy
           161 TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 220
Db
        805 TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGC 865
QУ
           221 TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 280
Db
        866 CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
Qу
               281 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 340
Db
        926 AAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
QY
                                     341 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 400
Db
        986 TCCGAGACTTGAAGAAAACGGCCC 1009
Qу
           401 TCAGGGATTTGAAGAGACCAGCTC 424
DU
RESULT 3
US-09-867-701-4953
Sequence 4953, Application US/09867701
 Patent No. US20020132237A1
 GEMERAL INFORMATION:
  MPPLICANT: Aglate, Paul A.
  APPLICANT: Jones, Robert
  APPLICANT: Harlocker, Susan L.
  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
  TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
  TELE REFERENCE: 210121.497
  CURRENT APPLICATION NUMBER: US/09/867,701
  SURRENT FILING DATE: 2001-05-29
  TUMBER OF SEQ ID NOS: 10912
  MOFTWARE: FastSEQ for Windows Version 4.0
 SEQ. ID NO 4953
   LENGTH: 450
   TYPE: DNA
   ORGANISM: Homo sapien
US-09-867-701-4953
                            Score 244.8; DB 10;
                                              Length 450;
                     24.1%;
 Query Match
                     77.3%; Pred. No. 5.3e-60;
 Best Local Similarity
                                                         Gaps .
                           0; Mismatches
                                         87;
                                             Indels
 Metches 297; Conservative
        626 TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
QУ
                  27 TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC 86
٦h
        636 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
Qу
            37 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATGACAAAATACA 146
Do
        745 TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
Qy.
            147 TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 206
Dh
```

```
806 TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGC 865
QУ
          207 TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 266
Db
       856 CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
Qу
          257 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 326
Db
       926 AAAAAGAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
Qy ...
          327 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 386
Db
       986 TCCGAGACTTGAAGAAAACGGCCC 1009
Qy
           387 TCAGGGATTTGAAGAGACCAGCTC 410
Db
RESULT 9
US-09-910-943-35
; Sequence 35, Application US/09910943
 Patent No. US20020081610A1
 GENERAL INFORMATION:
  APPLICANT: Hemmati-Brivanlou, Ali
  APPLICANT: Altman, Curtis
  THITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
  FILE REFERENCE: 7529/1G148US1
  CURRENT APPLICATION NUMBER: US/09/910,943
  CURRENT FILING DATE: 2001-07-23
  MUMBER OF SEQ ID NOS: 742
  SOFTWARE: PatentIn version 3.1
 SEO ID NO 35
   LENGTH: 762
   TYPE: DNA
   ORGANISM: Xenopus laevis
   FEATURE:
   NAME/KEY: misc feature
   LOCATION: (1)..(762)
   OTHER INFORMATION: n may be a or g or c or t/u
U3-09-910-943-35
                          Score 210.8; DB 9; Length 762;
                    20.8%;
  Query Match
                    78.6%; Pred. No. 4.8e-50;
  Best Local Similarity
                                                  3; Gaps
                                                           3;
  Matches 287; Conservative 0; Mismatches 75;
                                          Indels
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Q_{\mathcal{F}}
           397 ATGAAGAAAATGCCATTGTTCAGCAAGTCACATAAAAATCCGGCTGAGATTGTTAAAACT 456
Db
         Qу
           457 CTGAAGGACAACATGGCCCTGCTGGAAAGGCAGGACAAAAAAACTGAAAAGGCCTCTGAA 516
D:x
        Qy
           517 GAAGTGTCTAAATCTCTTCAAGCTACAAAAGAGATTTTGTGTGGGACAGGGGACAAAGAA 576
Db
```

```
181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
QУ
            577 CCTCAGACAGAGACGGTGGCTCAGCTCGCACAAGAACTGTACAACAGTGGCTTGTTGGTT 636
ďQ
         241 ACACTGATAGCTGACC-TGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGAT 299
QУ
            637 ACTTTAATAGCCCACCTTGCATCTCATAGATTTTGANGGCAAGAAAGATGTATCTCAGAT 696
Dio
         300 ATTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAG 359
ΩУ
            697 ATTCNAC-ACATCCTGAGAAAACAGATTGGCACTCGGAGTNC-CCTGTGGAGTATATCAA 754
Db
         350 TGCTC 364
             1 1 1
         755 TTCCC 759
RESULT 10
US-09-954-456-1453/c
; Sequence 4453, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
  APPLICANT: Young, Paul
   TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents
Using Cancer Gene
 FINITEDE OF INVENTION: Sets
   NULL REFERENCE: 689290-76
   CURRENT APPLICATION NUMBER: US/09/954,456
   CURRENT FILING DATE: 2001-09-18
   PRIOR APPLICATION NUMBER: US/60/233,617
   MRIOR FILING DATE: 2000-09-18
   PRIOR APPLICATION NUMBER: 18/60/234,052
   PRIOR FILING DATE: 2000-09-20
   PRIOR APPLICATION NUMBER: US/60/234,923
  PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
   PRIOR FILING DATE: 2000-09-25
   PRIOR APPLICATION NUMBER: US/60/235,637
   PRIOR FILING DATE: 2000-09-26
   PRIOR APPLICATION NUMBER: US/60/235,638
   PRIOR FILING DATE: 2000-09-26
   PRIOR APPLICATION NUMBER: US/60/235,711
   PRICE FILING DATE: 2000-09-27
   PRIOR APPLICATION NUMBER: US/60/235,720
   PRIOR FILING DATE: 2000-09-27
  PRIOR APPLICATION NUMBER: US/60/235,840
   PRIOR FILING DATE: 2000-09-27
   PRIOR APPLICATION NUMBER: US/60/235,863
   PRIOR FILING DATE: 2000-09-27
   NUMBER OF SEQ ID NOS: 2276
   SOFTWARE: PatentIn version 3.0
  SEO ID NO 1453
    LENGTH: 387
    TYPE: DNA
    ORGANISM: Homo sapiens
```

US-09-954-456-1453

```
19.2%; Score 195; DB 10; Length 387;
 Ouery Match
                   100.0%; Pred. No. 1.1e-45;
 Best Local Similarity
                                                     Gaps
                                                  0:
                       0; Mismatches
                                       0;
                                          Indels
 Matches 195; Conservative
       820 TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 879
ÖА
          387 TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 328
Db
       Qγ
           Db
       940 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 999
           \Omega \lambda
       267 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 208
Db
       1000 AAAACGGCCCCTTGA 1014
Q7
           20% AAAACGGCCCCTTGA 193
D^{3}
RESULT 11
US-09-530-107-481/c
; Sequence 481, Application US/09880107
 Patient No. US20020142981A1
 CLUBERAL INFORMATION:
  MARRICANT: Horne, Darci T.
  PARATCANT: Vockley, Joseph G.
  APPLICANT: Scherf, Uwe
  APPLICANT: Gene Logic, Inc.
  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
  FILE REFERENCE: 44921-5028-WO WAS 200
  CURRENT APPLICATION NUMBER: US/09/880,107
  CURRENT FILING DATE: 2001-06-14
  FRIOR APPLICATION NUMBER: US 60/211,379
  PRIOR FILING DATE: 2000-06-14
   PRIOR APPLICATION NUMBER: US 60/237,054
   PRIOR FILING DATE: 2000-10-02
  NUMBER OF SEQ ID NOS: 3950
   SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 481
   ENGTH: 387
   TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
    OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA234362
US-09-380-107-481
                    . 19.2%; Score 195; DB 10; Length 387;
  Query Match
  Pest Local Similarity 100.0%; Pred. No. 1.1e-45;
                                                       Gaps
                          0; Mismatches
                                            Indels
                                        0;
  Matches 195; Conservative
        820 TTTCATGTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 879
            QΥ
        337 TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 328
 Db
         Qу
```

```
Db
        940 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 999
Qу
            267 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 208
Db
       1.000 AAAACGGCCCCTTGA 1014
Qy
            207 AAAACGGCCCCTTGA 193
Db
RESULT 12
US-10-257-826A-113
; Sequence 113, Application US/10257826A
 Publication No. US20030181407A1
; GENERAL INFORMATION:
  APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
  APPLICANT: PALIN, Marie-France
  APPLICANT: POMAR, Candido
  APPLICANT: GARIEPY, Claude
  TITLE OF INVENTION: Steatosis-modulating factors and uses
  TITLE OF INVENTION: thereof
  FILE REFERENCE: 14654 2US
   CURRENT APPLICATION NUMBER: US/10/257,826A
  CURRENT FILING DATE: 2002-10-17
   PRIOR APPLICATION NUMBER: 60/197936
   PRIOR FILING DATE: 2000-04-17
   FRIOR APPLICATION NUMBER: PCT/CA01/00509
  TRRICR FILING DATE: 2001-04-12
                                                           dis
   MARKER OF SEQ ID NOS: 305
                                                           300
   SOFTWARE: FastSEQ for Windows Version 4.0
  and ID NO 118
    NGTH: 722
    YPE: DNA
    CRGANISM: Artificial Sequence
    OTHER INFORMATION: Artificial sequence
    OTHER EMFORMATION: Muscular steatosis
    OTHER INFORMATION: Porcine
    FEATURE:
    NAME/KEY: misc feature
    LOCATION: (1) ... (722)
    OTHER INFORMATION: n = A, T, C or G
US-10-257-826A-118
                      16.7%; Score 169.8; DB 13; Length 722;
  Query Match
                      60.1%; Pred. No. 3.1e-38;
  Best Local Similarity
                            0; Mismatches 196;
                                               Indels
  Matches 303; Conservative
         347 TGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCC 406
 ΩУ
                        2 TGGTGAATNCCTCTGCCCCCACNGAATTTTTGGTCATGGTANTNGAAGGGGATNAATNTT 67
 Db
         407 CACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTG 466
 \Omega V
             68 CCGAAATTNCNNTTAATTGGGGNATNATGGTNAGANAATGCCTTNGACCTNNACCGCTTG 127
 Db
```

```
467 CCAAAATCATCCTCTTTTC--TAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTC 524
QУ
           128 CCAAAATCATTTTGNGGGCCGAACACAGTTTATAGAGATCTTCACATATGTCTAAATGTN 187
Db
        525 AACA-TTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATA 583
С'n
           138 ANCATTTTNACATATCTTTACATNCNNTTNCCNCATTTTNNGNNTTACTTTCACGACATA 247
ďŒ
        584 AAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGA 643
Qу
            248 TATTGCTCACNGCGCAANTTTTGGAACANCATTATGATANATTTTTCAGTGAATATGATG 307
Db
        644 AATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGC 703
Q7
           308 AAGNGCTTCATTCTTAAAATTATGTGGCCACAAGACAATCACTGAAGCTTCTCGGNGAAC 367
Db
        704 TGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACC 763
Qν
           368 TACTACTANATAGACNCNACTTCNCCANTATGACCACATACCTCATTAAACCTGNGNACC 427
Db
        764 TGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAA-CATCCAGTTTGAAGCCTTT 822
Qу
           488 T-CCATTAATGATGAACCTGCCTGCAGACAAAAGTCGGAACCTTCCANTTTGAGGGCTTN 486.
Ob
       . 823 CATGTTTTAAGGTGTTTGTGGCC 846
्रा
            487 CACGTTTTTAANGGGGNTGTNNNC 510
MESSELF 13 -
US-15-257-326A-119
; Sequence 119, Application US/10257826A
: Publication No. US20030181407A1
; GENERAL INFORMATION:
 APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
  APPLICANT: PALIN, Marie-France
   APPLICANT: POMAR, Candido
   APPLICANT: GARIEPY, Claude
   TITLE OF INVENTION: Steatosis-modulating factors and uses
   TITLE OF INVENTION: thereof
  FILE REFERENCE: 14654-2US
  CURRENT APPLICATION NUMBER: US/10/257,326A
   CURRENT FILING DATE: 2002-10-17
   FRIOR APPLICATION NUMBER: 60/197936
   PRIOR FILING DATE: 2000-04-17
   PRIOR APPLICATION NUMBER: PCT/CA01/00509
   PRIOR FILING DATE: 2001-04-12
   NUMBER OF SEQ ID NOS: 305
   SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 119
    LENGTH: 700
   TYPE: DNA
    ORGANISM: Artificial Sequence
```

FEATURE:

OTHER INFORMATION: Artificial sequence OTHER INFORMATION: Muscular steatosis

```
OTHER INFORMATION: Porcine
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (1)...(700)
  OTHER INFORMATION: n = A, T, C or G
US-10-257-826A-119
                   16.4%; Score 166.6; DB 13; Length 700;
 Query Match
                   60.1%; Pred. No. 2.6e-37;
 Best Local Similarity
                         0; Mismatches 197; Indels
                                                 5; Gaps
 Matches 304; Conservative
       345 TGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGC 404
QУ
                           1111
         5 TCTGGTGAATCCCTCTGCCCCCACNGAATTTTTGGTCATGGTANTNGAAGGGGATNAATN 65
Db
       405 CCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACT 464
            QΥ
        66 TTCCGAAATTTCGATTAATTGGGGNATNATGGTNAGANAATGCCTTNGACCTCCACCGCT 125
Db
       465 TGCCAAAATCATCCTCTTTTC--TAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTG 522
           QУ.
       126 TGCCAAAATCATTTTGNGGGCCGAACACAGTTTATAGAGATCTTCACATATGTCTAAATG 185
Db
        TCAACA-TTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACA 581
Qγ
           245:
        TAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGA 6418
           -1111
        24.6 TATATTGCTCACNGCGCAANTTTTGGAACANCATTATGATANATTTTTCAGTGAATATGA 305.4
115
        42 GAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGA 701
            365 TGAAGNGCTTCATTCTTAAAATTATGTGGCCACAAGACAATCACTGAAGCTTCTCGGNGA 365
图图
        702 GCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAA 761
Q7
            356 ACTACTACTANATAGACNCNACTTCNCCANTATGACCACATACCTCATTAAACCTGNGNA 425
70 h
        762 CCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAA-CATCCAGTTTGAAGCCT 820
               Сy
        426 CCT-CCATTAATGATGAACCTGCCTGCAGACAAAAGTCGGAACCTTCCANTTTGAGGGCT 484
Db
        821 TTCATCTTTTTAAGGTGTTTTGTGGCC 846
ΩУ
           485 TNCACGTTTTTAANGGGGNTGTNNNC 510
Db
RESULT 14
 US-09-770-445-592
 ; Sequence 592, Application US/09770445
  Patent No. US20020023281A1
 ; GENERAL INFORMATION:
   APPLICANT: Gorlach, Jorn
```

APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.

```
APPLICANT: Raines, Tracy M.
 APPLICANT:
           Yu, Yang
            Rameaka, Joshua G.
 APPLICANT:
            Page, Amy
  APPLICANT:
            Matthew, Abraham V.
  APPLICANT:
            Ledford, Brooke L.
  APPLICANT:
            Woessner, Jeffrey P.
  APPLICANT:
  APPLICANT: Haas, William David
  APPLICANT: Garcia, Carlos A.
  APPLICANT: Kricker, Maja
            Slader, Ted
  APPLICANT:
            Davis, Keith R.
  APPLICANT:
            Allen, Keith
  APPLICANT:
            Hoffman, Neil
  APPLICANT:
            Hurban, Patrick
  APPLICANT:
  TITLE OF INVENTION: Expressed Sequences of Arabidopsis
  TITLE OF INVENTION: thaliana
  FILE REFERENCE: 2023US (PARA-012PRV)
  CURRENT APPLICATION NUMBER: US/09/770,445
  CURRENT FILING DATE: 2001-01-26
  PRIOR APPLICATION NUMBER: US 60/178,472
  PRIOR FILING DATE: 2000-01-27
  NUMBER OF SEQ ID NOS: 999
  SOFTWARE: FastSEQ for Windows Version 4.0
 CEQ ID NO 592
   LENGTH: 861
   TYPE: DNA
  OFGANISM: Arabidopsis thaliana
97-69-770-445-592
                      15.4%; Score 156; DB 9; Length 861;
 opery March
                      55.8%; Pred. No. 3.4e-34;
Best Local Similarity
                            0; Mismatches 250; Indels
                                                          Gaps
 Matches 319; Conservative
        394 GGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGA 453
            12 GGGTTTGAAAACACCGATATGGCGTTACACTATGGTACTATGTTTAGAGAGTGCATCCGT 71
מכי
        454 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 513
Q_{J}^{*}
                    72 CATCAGATTGTTGCAAAATATGTTTTGGACTCGGAGCACGTGAAGAAGTTTTTTTACTAC 131
Db
        514 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 573
07
                                            132 ATACAGCTTCCCAATTTCGACATTGCTGCTGATGCTGCTGCAACTTTTAAGGAACTTCTG 191
777
        574 ACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 633
Qy
                           11 11 11 11
        192 ACAAGGCACAAGTCTACAGTTGCTGAGTTTCTCATTAAGAATGAAGACTGGTTTTTTGCA 251
D'a
        634 GACTATGA---GAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAG 690
QУ
                                                252 GACTACAACTCAAAGCTTCTTGAATCAACTAATTATTACCCGACGGCAAGCTATTAAG 311
Db
        691 CTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGC 750
QУ
             312 TTGTTGGGCGATATATTATTGGATAGGTCAAATTCGGCTGTGATGACGAAGTATGTGAGC 371
Db
```

```
751 AAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAG 810
             QУ
        Db
        811 TTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTATT 870
            QУ
        432 ATAGAAGCTTTCCATGTTTTCAAGCTGTTTGTAGCGAACCAAAACAAGCCTTCAGACATC 491
Db
        871 GTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAA 930
Qу
           492 GCCAACATTCTGGTGGCAAACAGAACAGCTTCTGAGATTGTTGGCTGATATCAAGCCG 551
Db
        931 GAAAGGACGGATGATGAGCAGTTCGCTGACGA 962
QУ
            552 GACAAAGAGGACGAGAGGTTTGACGCAGACAA 583
RESULT 15
US-09-923-876-1251
; Sequence 1251, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
  NOPLICANT: Lalgudi, Raghunath V.
  APPLICANT: Kamigaki, Laura Y. (Ito)
  DEPLICANT: Sherman, Bradley K.
  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN
 TIME
                                                          33.3.1.2
  COMME REFERENCE: PL-0012-1 CON
  MEDIERENT APPLICATION NUMBER: US/09/923,876
  QUERENT FILING DATE: 2001-08-06
  OR APPLICATION NUMBER: 09/298,329
   CREOR FILING DATE: 1999-04-21
   FRIOR APPLICATION NUMBER: 60/085,331
  PRIOR FILING DATE: 1998-05-05
  NUMBER OF SEQ ID NOS: 6332
  SOFTWARE: PERL Program
  SEC ID NO 1251
   ICNCTH: 262
    TYPE: DNA
   ORCANISM: Zea mays
    FEATURE:
   MAME/KEY: misc_feature
   OTHER INFORMATION: Incyte ID No. US20020013958A1 700158378H1
    MAME/KEY: unsure
  LOCATION: 148
 , CTHEE INFORMATION: a, t, c, g, or other
 US-09-923-876-1251
                       7.3%; Score 74.2; DB 9; Length 262;
  Quary Match
  Best Local Similarity 55.5%; Pred. No. 6.4e-11;
  Matches 142; Conservative 0; Mismatches 114; Indels
         311 TCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTC 370
 Qv
             1 111111 11 11 1 1 1 1 1111111 1111 1
           7 TGTTGAGACAGAGGTTGATGAAAGCTATTGTTGCGTCCAGTATATTGAAAATCATTTTG 66
 Db
```

Qy Db		ATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGA 430
~1		TTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATC 490
Db	127	ACATGTTGCGAGAATGCATAANATATCCTACACTTGCAAAATATATTTGGAGTCAAGCA 186
Q.Y		AATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCT 550
Db	187	GCTTCGAGTTTTTCCAGTATGTTGAATTGTCAAACTTCGATATTGCATCTGATGCTC 246
Qy		TTGCTACTTTCAAGGA 566
Db	247	TGAACACTTTCAAGGA 262

Search completed: January 6, 2004, 05:04:45
Job time: 1400 secs